

Cleft Constructions in Discourse

J. L. Delin

PhD

University of Edinburgh

1989



Declaration

I declare that this thesis has been composed by myself, and the research reported herein has been conducted by myself unless otherwise indicated.

J. L. Delin
Edinburgh, September 28th 1989

For my Parents

for showing me that language was fun to play with

Acknowledgements

There are many people who have guided, helped and influenced me during the writing of this thesis. I would like to thank them all individually, but will be content in this acknowledgement to single out an outstanding few. First of all, I would like to thank my supervisors, Ewan Klein and Graeme Ritchie, for their willingness to listen to and discuss my ideas, curb my worst intellectual excesses, and inject a little formal rigour here and there. I have learnt many important lessons from them, not least the value of choosing a story, and sticking to it.

Of the many others who have influenced this work, I would like in particular to record my gratitude to Bob Ladd, whose refreshing perspective at a critical stage in the research provided vital encouragement. Others who have been willing to discuss individual issues, and provide constructive criticism of various parts of this thesis, include Steven Bird, Ronnie Cann, Robert Dale, Marc Moens, Jon Oberlander, Rob van der Sandt, Antonio Sanfilippo, and Pete Whitelock: their efforts should not go unrecorded.

There are many people whose presence at home or at work have made things easier, more fun, or both. Helen Pain in particular deserves my thanks not only for providing me with a job during the last two years of this research, but for going out of her way to ensure the best conditions for the thesis to be finished. Other special friends who have helped me along the way are Kathleen King, Joe Levy, Mimo Caenepeel, Karen Valley (with special thanks for her last-minute sprint at the keyboard), Lesley Nethery, Jon and Vina Oberlander, Cath Ardin, Adam Gardner, and especially Jeremy Olsen for trips in the car, cat-minding, and generally Being There. I would also like to thank my friends and classmates at the Centre for Cognitive Science for making it a good place to work.

Thanks should also go to some prominent figures who were influential early on in the process leading to this thesis. Clive Williams noticed me in junior school and helped me get onto the right path. Mrs. Eason taught me not only how to make baked egg custard but a lot about methodology, including the importance of washing up as I went along. Sister Mary Angela instructed me in Italian and principles for living. Mike Stubbs, Margaret Berry, and Chris Butler from my Nottingham days got me excited about linguistics. I am grateful to all of them.

Most important in my past and present are my family, Ma, Pa, Annie, Sue, and Tara, who have been a solid base all along for everything I do. Special thanks are due to Annie, who Olive Schreiner has described better than I ever could.

Finally, love and thanks to my Edinburgh family, Robert and William, who have each in their own way left their marks on these pages. Now that this is over, I'm looking forward to being better company to them both.

*'It is this, it is this—' 'We have had that before!'
The Bellman indignantly said.
And the Baker replied, 'Let me say it once more.
It is this, it is this that I dread!'*

Lewis Carroll, *The Hunting of the Snark*
cited in Pullum [1989]

Abstract

This thesis presents an analysis of the structure and function of cleft constructions in discourse. Drawing on a corpus of naturally-occurring spoken and written data, we present a multi-layered explanation of how *it*-clefts, *wh*-clefts, and reverse *wh*-clefts are different from non-clefts, and from one another. After a review of previous research on clefts in discourse, we explore the aspects of syntax, semantics, and pragmatics relevant to the structure and function of all three types of cleft. The discussion falls into three main parts:

- An analysis of the three cleft types, within the framework of Generalized Phrase Structure Grammar (cf. Gazdar *et al.* [1985]), in which particular attention is paid to the variety of constituents that can appear in particular positions in each type. The output of the grammar rules is compared to the examples that occur in the corpus of data.
- A treatment of cleft presupposition in terms of an analogy (suggested by van der Sandt [1988]) between presupposition and the treatment of pronominal anaphora in Kamp's [1981] Discourse Representation Theory.
- An examination of the range of accentual patterns, presuppositional relations, and information structures typically appearing in clefts of all three kinds. We show that marked distinctions exist between the three cleft types in terms of all these factors, and suggest ways in which this helps to differentiate the range of discourse contexts in which clefts in general, and each cleft type in particular, are appropriate.

At the end of the thesis we point to an analogy between the formal model for clefts presented and a psychological model of sentence processing. We also suggest how the conclusions regarding both the structure and function of clefts as a class of construction and the distinction between the three types of cleft could be synthesised in a decision procedure for syntactic choice. Finally, we suggest some related areas for further research.

Contents

1	Introduction	1
1.1	Background to the Research	1
1.2	Preliminaries	4
1.2.1	The Aims of the Research	4
1.2.2	Starting Points	4
1.3	The Object of Study	8
1.4	The Structure of the Rest of the Thesis	9
2	An Overview of Previous Research	12
2.1	Accounts Based on Focus-Marking	13
2.1.1	Chafe	14
2.1.2	Creider	17
2.1.3	Quirk <i>et al</i>	22
2.1.4	Two Theories of Focus in Computational Linguistics	24
2.1.5	Geluykens	28
2.1.6	Clefts and Focus-Marking: Conclusion	35
2.2	Accounts Based on the Marking of Other Information Types	37
2.2.1	Prince	37
2.2.2	Lapolla	42

2.2.3	Werth	45
2.2.4	Dividing Given from New Information	50
2.2.5	Clefts and the Marking of Other Information Types: Conclusion	52
2.3	An Overview of Previous Research: Summary	53
3	The Syntax of Cleft Constructions	55
3.1	The <i>It</i> -Cleft	56
3.1.1	Phrase-Structure Rules for <i>it</i> -clefts	57
3.1.2	<i>It</i> -clefts and Topicalisation	65
3.1.3	Agreement in <i>It</i> -clefts	68
3.1.4	Summary	70
3.2	The <i>Wh</i> -Cleft	70
3.2.1	Phrase-Structure Rules for <i>Wh</i> -clefts	71
3.2.2	The Instantiation of XP	73
3.2.3	The Analysis of <i>Wh</i> -clefts	78
3.2.4	<i>Wh</i> -clefts and other <i>Wh</i> -Constructions	81
3.2.5	Agreement in <i>Wh</i> -clefts	89
3.2.6	Summary	91
3.3	The Reverse <i>Wh</i> -Cleft	92
3.3.1	Phrase-Structure Rules for Reverse <i>Wh</i> -clefts	92
3.3.2	The Instantiation of XP	95
3.3.3	Summary	99
3.4	A Note on the Corpus Data	100
3.5	The Syntax of Cleft Constructions: Summary	102
4	Aspects of Cleft Semantics	106

4.1	Cleft Constructions and Referentiality	109
4.1.1	Higgins' View of <i>Wh</i> -clefts	110
4.1.2	The Referential Relations of Clefts	113
4.2	Cleft Constructions and Presupposition	115
4.2.1	Characterising Presupposed Information	119
4.2.2	Accounts of Cleft Presupposition	123
4.2.3	Requirements for a Theory of Cleft Presupposition	138
4.2.4	A Model for Presupposition	141
4.3	Cleft Constructions and Uniqueness	147
4.3.1	The Evidence for Uniqueness	148
4.3.2	The Analogy with Definiteness	152
4.4	Clefts in Discourse Representation Theory	153
4.4.1	A Brief Introduction to DRT	154
4.4.2	The Treatment of Pronominal Anaphors	155
4.4.3	The Treatment of Presuppositions	157
4.4.4	Cleft Presuppositions as Anaphors	160
4.4.5	Clefts and Maximality	161
4.4.6	The Projection Problem and DRT	163
4.5	Aspects of Cleft Semantics: Summary	165
5	Pragmatic Aspects of Cleft Constructions	166
5.1	The Accentual Patterns of Clefts	168
5.1.1	Five Types of Accentual Pattern	169
5.1.2	Profile of Tone Groups in the Data	173
5.2	Clefts and the Serial Ordering of Information	174

5.3	Cleft Presuppositions in Discourse	178
5.3.1	A Rhetorical Constraint on Presupposition	180
5.3.2	Three Types of Presuppositional Relation	182
5.3.3	The Distribution of Presuppositional Relations	196
5.3.4	Clefts and Presupposition: Summary	203
5.4	The Information Structure of Clefts	203
5.4.1	Analysing Information Structure	205
5.4.2	General Features of Cleft Information Structure	212
5.4.3	Features of Information Structure Specific to Cleft Types	216
5.5	Information Structure and Presupposition: A Synthesis	222
5.5.1	The Relative Roles of Presupposition and Information Structure	222
5.5.2	Why Presuppose?	223
5.6	Pragmatic Aspects of Cleft Constructions: Summary	228
6	Conclusions and Further Research	230
6.1	The Claims Made in this Thesis	231
6.2	Choosing Clefts in Context	236
6.3	Directions for Further Research	239

Chapter 1

Introduction

This thesis is about CLEFT CONSTRUCTIONS, particularly the factors that might determine a speaker's or writer's decision to use such a construction in a particular discourse context. The types of cleft-construction that will be considered are IT-CLEFTS (sometimes termed CLEFTS), WH-CLEFTS (sometimes termed PSEUDO-CLEFTS), and REVERSE WH-CLEFTS (or INVERTED PSEUDO-CLEFTS). Examples of each respectively appear below:

- (1-1) a It was Karen who was doing the typing.
 b What I'm interested in is aeroplane modelling.
 c That's what I meant.

The purpose of the thesis is to examine a range of syntactic, semantic and pragmatic-contextual factors that are relevant in differentiating clefts from other sentence types and from one another, and to present an account of their functions in connected discourse in terms of a synthesis of these factors.

1.1 Background to the Research

Cleft constructions of all three types first began to attract attention in the 1970's (cf. Akmajian [1970], Harries [1972], Hankamer [1975], Gundel [1975, 1977], Halvorsen [1978], Higgins [1979] *inter alia*) where the majority of research was couched in the terms of Transformational Grammar. For this reason, the discussion centered upon theory-internal debates concerning which kinds of cleft were base-generated, and which were

transformationally driven, and whether the transformations involved were pre- or post-cyclic.

Although it is possible to extract some concrete observations with respect to cleft syntax from this literature, the transformational tradition left two particular legacies with respect to the study of clefts. One, perhaps appropriate given the syntactic nature of the research involved, was a methodology almost entirely based on the study of decontextualised examples, judgements regarding the grammaticality or acceptability of the data involved being left to the discretion of the researcher. This had long-term effects on research on clefts in general, not just in the field of syntax, in that the assumptions made about the syntactic, prosodic and functional variety that existed in the data seemed to underestimate, often drastically, what corpus-based study can actually reveal to be the case.

The second legacy, which stems from the first, is the assumption that *it*-clefts and *wh*-clefts, as truth-functionally equivalent constructions, are interchangeable in context. Akmajian states this explicitly:

There is a great deal of similarity between cleft and pseudo-cleft sentences.
For example, the following two sentences:

- (1-2) a The one who Nixon chose was Agnew.
b It was Agnew who Nixon chose.

are synonymous, share the same presuppositions, answer the same questions,
and in general they can be used interchangeably.

[Akmajian 1970:149]

Eight years after Akmajian's claim, Prince [1978] produced an important rejoinder: she argued in her paper *A comparison of wh-clefts and it-clefts in discourse* that the two construction types were distinct in structure, and in discourse function, supporting her arguments with naturally-occurring, contextualised examples. Prince was able to show not only distinct functions for *it*-clefts and *wh*-clefts, but the existence of a heterogeneity *within* the two cleft types that was an entirely new observation. In many ways, the research reported in this thesis takes this important paper of Prince's as a starting point.

As with all research, Prince's work left many questions unanswered. At what level of description could the distinctions between cleft constructions be differentiated from non-clefts? Was it really the case that *it*-clefts and *wh*-clefts served to convey entirely distinct types of information, in all cases? What was the true extent of the heterogeneity Prince claimed regarding *it*-clefts in particular, and had she even underestimated it? In particular, what could be said for reverse *wh*-clefts, which had rarely received any attention, either in Prince's work or in the work that led up to it?

Apart from the important stimulus of Prince's paper, a second influence in this research has been the parallel work on cleft semantics that has been going on often independently of research into their syntax and discourse function. Many researchers (notable among them being Chomsky [1971], Keenan [1971], Stalnaker [1974], and Gazdar [1979]) have suggested a role for clefts as PRESUPPOSITIONAL constructions, on a variety of definitions of presupposition. Although it was generally agreed that clefts do convey presuppositions, very little research had been done into how presupposition might relate to discourse notions such as information structure and focus. For example, how could a syntax-based view of presupposition—where a presupposed proposition was postulated on the basis of the syntactic structure of the sentence—be related to accounts of cleft constructions that appeal to notions of what is 'Given' or 'New' in context? What did it mean for 'New' information to be presented as part of the presupposition?

Both kinds of account clearly have their advantages for the description of clefts. The accounts based on information structure have their strength in the fact that they address how an utterance relates to the current state of the discourse context, and the shared knowledge and assumptions of speaker and hearer. The weakness of such accounts, however, is that they have no power to explain the existence of clefts as a particular syntactic structure: on the basis of information structure alone, clefts of various types are often equivalent to other structures. The strength of syntax-based accounts of cleft presupposition, on the other hand, is that they point to the syntactic form of the cleft having a recognisable function that distinguishes it from many other sentence types: the fact that it signals presupposition. The problem with such accounts, however, lies in the problem of relating presupposition to what is currently going on in the discourse in which such presupposing constructions appear, and what function the presupposition serves. If no such relationship can be found, the notion of presupposition, however

neatly derived from the surface form, is merely a sterile artefact.

Much of this thesis centres around reconciling these two accounts, and synthesising the valuable parts of each into an account of clefts that has a formal basis but is able to describe and account for the wide range of types and functions of cleft that appear in a corpus of data. Some of the work described here is consolidatory, formalising aspects of cleft structure—particularly information structure—about which observations have been made before. Other parts of the thesis are analytical, exploring and exposing the range of data to be described. Perhaps the most important part of the thesis consists of a synthesis of the syntactic, semantic and pragmatic facts about clefts into a multi-layered account of the distinctions between the cleft types, and the distinctions between cleft and non-cleft sentences.

1.2 Preliminaries

1.2.1 The Aims of the Research

On the basis of the discussion above, the aims of the research reported here can be summarised in the following points.

- To expose the syntactic, prosodic and pragmatic heterogeneity of the cleft data on the basis of a corpus of naturally-occurring written and spoken data.
- To provide a formally-based but largely theory-independent means of describing the data.
- To motivate the distinctions between cleft and non-cleft sentences, and between the three types of cleft, at a variety of levels of description.
- To show how factors identified at these various levels can be synthesised into an account of the structure and function of clefts in discourse.

1.2.2 Starting Points

In the research to be reported, we begin with a number of background assumptions.

The Principles of Utterance Design

We assume that in order to convey a message in a particular context, speakers and writers have at their disposal a range of message-carrying devices, i.e., different types of sentence. Many of these sentence types are truth-functionally equivalent, and the choice between them has to be made on the basis of a variety of syntactic, semantic and pragmatic factors. It is assumed here that sentence type is chosen on the basis of the following aims on the part of the speaker or writer:

- to mark the message being conveyed with particular information regarding its relationship with the discourse context; and
- to observe as far as possible co-operative principles ensuring the comprehensibility of the message.

While the research reported is not couched in procedural terms—it does not constitute an algorithm for choosing cleft constructions, for example—it is assumed here that a complete explanation of clefts in discourse is one that contains sufficient descriptive detail to discriminate between clefts and non-clefts and between the three cleft types, in the majority of cases.

A Corpus-Based Study

In the description of early work on cleft syntax given above, it was noted that much underestimation of the heterogeneity of the clefts data resulted from the study of idealised, decontextualised examples. The research reported here is predominantly corpus-based, drawing on a sizeable corpus of data in the description of the syntactic and pragmatic variety of clefts in particular.

The data used throughout this study are included in the Appendix to this thesis, and are as follows:

- A corpus of spoken data, taken from the Survey of English Usage corpus (Svartvik [1980]), consisting of 54 *it*-clefts, 63 *wh*-clefts and 106 reverse *wh*-clefts.

- A corpus of written data, taken from the Lancaster-Oslo-Begen (LOB) on-line corpus, consisting of 165 *it*-clefts, 72 *wh*-clefts and 38 reverse *wh*-clefts.
- My own collection of casual conversational data, consisting of 37 *it*-clefts, 36 *wh*-clefts and 16 reverse *wh*-clefts.

The total tokens amount to 256 *it*-clefts, 171 *wh*-clefts and 160 reverse *wh*-clefts.

In one case, the corpora studied for the purposes of more detailed discussion were cut down in size: in the discussion of prosody and pragmatic aspects of the cleft, two sub-corpora of 150 examples each, drawn from the spoken data, were used in the analysis. This is made clear where it occurs, and the particular examples chosen are identified in the appendix.

On occasion, where conclusions are to be drawn on the basis of the frequency of a particular phenomenon appearing in the data, Chi-squared tests are performed to verify the reliability of distribution in the sample. This is made clear where relevant, and the degree of significance of the test pointed out in each case.

Constructed data appear in cases where particular points of grammaticality or acceptability are at issue, particularly for the purposes of the discussion of cleft syntax in chapter 3.

A Multi-Layer Approach

As we noted above, the motivation for the use of clefts in discourse is shown in the thesis not to reside at a single level of description. Instead, it is demonstrated that aspects of syntax, semantics and pragmatics are all involved in carving out a functional niche for cleft constructions as a group, and for the three individual cleft types in particular. In the discussion of previous research given in chapter 2, we point out that accounts that attempt to explain cleft function at a single level—in terms of information ordering, for example—can never be sufficiently general, and in almost all cases fail to capture crucial factors determining the function of clefts that reside at other levels. Accordingly, the core of the thesis consists of a discussion of cleft syntax, semantics and pragmatics respectively. In addition, in the conclusion to the discussion of pragmatic aspects, we

suggest a model of the relative roles of presupposition and information structure in determining the function of clefts, and in the concluding chapter of the thesis we make some suggestions regarding how the conclusions reached could interact in a decision procedure for choosing clefts in context.

Terms of Reference

At the beginning of this chapter, we identified the three cleft types using the terms IT-CLEFT, WH-CLEFT and REVERSE WH-CLEFT. These are the terms used throughout the thesis. They are adopted in preference to CLEFT, PSEUDO-CLEFT and INVERTED PSEUDO-CLEFT, partly because the latter terms carry transformational overtones not in keeping with the approach adopted here, and partly because people seem to have a certain amount of difficulty remembering which term applies to which cleft type. Often in the thesis we refer to all three kinds of construction as CLEFT CONSTRUCTIONS, or sometimes simply as CLEFTS. As far as the discussion of the various parts of cleft constructions is concerned, we will avoid the terms FOCUS and PRESUPPOSITION as often used to describe particular syntactic constituents. This is because such terms seem to indicate a blurring of the distinction between syntactic form and pragmatic function, and to adopt them would therefore pre-empt a great deal of the discussion contained in this thesis. The terms we will adopt instead are therefore as follows. For clefts of all three types, the immediate complement of the copula—for example, *the box* in the examples in (1-3)—will be termed the cleft's HEAD. The embedded clausal constituent in *it*-clefts will be termed the *it*-cleft COMPLEMENT, while in the case of *wh*-clefts and reverse *wh*-clefts the corresponding constituent is termed the WH-CLAUSE:

- (1-3) a It is the box that I'm standing on.
 b What I'm standing on is the box.
 c The box is what I'm standing on.

Other terms and forms of notation will be introduced where appropriate in the chapters to follow.

In talking about syntactic constructions that seem to be a departure from some 'canonical form', it is often easy to lapse into terminology that suggests that such constructions have undergone some transformational process in order to reach their current form. In this thesis, while we have tried to avoid the use of such terms, they may creep in from time to time simply because the transformationalists left behind them such a wealth of useful terms that at times it seems obtuse not to use them. However, the view taken here of 'marked' syntactic constructions such as clefts is not that they are transformed from some other surface form. Instead, clefts, canonical declaratives, and any other sentence types are seen as a range of possible realisations for a given propositional content. Declaratives are seen as distinct from 'marked' syntactic structures only in that they appear to be more neutral in terms of the markings that they convey in addition to their propositional content, which makes them a possible choice for speakers in a wider variety of contexts.

1.3 The Object of Study

There exist several categories of sentence type that are similar to the sentences which are the object of study in this thesis, some of which are also classed as cleft constructions. Before continuing, we should make clear which types fall within the scope of this research. As far as *it*-clefts are concerned, there is one other, rather rare construction that appears to be cleft-like, but is not studied here. The construction in question is exemplified in (1-4):

- (1-4) a It's a wise boy who knows his own father.
b It's an ill wind that blows nobody any good.

This sentence type is poorly understood, but appears to be distinct semantically at least from the *it*-cleft proper.

The *wh*-clefts and reverse *wh*-clefts studied here might be termed a *subset* of the class of pseudo-clefts, since we will not be concerned with some particular cases. Sentence types we will not address are:

- *Wh*-clefts and reverse *wh*-clefts whose clausal constituent modifies *the one/the ones*, such as *the one I'm looking for is John* and *my sister is the one that minds*.
- *Wh*-clefts and reverse *wh*-clefts with *all* in place of the *wh*-initiator of the clausal constituent, such as *all I want is a good night's sleep* and *a bran muffin was all I asked for*.

The exclusion of the first type above is simply due to the fact that cleft sentences with *the one* are one step further along a continuum of equative copular sentences, and seem to be the thin end of the wedge as far as the inclusion of all equatives of this type is concerned. To appreciate this, consider the progression from the *wh*-cleft in (1-5a) to the copular equative in (1-5d):

- (1-5) a What I want is a new hoover.
 b The one I'm looking for is John.
 c The thing that annoys me is their lack of charity.
 d The girl that spoke to you just now is his daughter.

Prince [1978] notes semantic distinctions between *wh*-clefts such as (1-5a) and (1-5b), and on this basis, confining the study to the 'headless' *wh*-cleft such as (1-5a) seems a coherent position to take. Semantic distinctions also exist between the *wh*-clefts proper and sentences with *all* as initiator of the clausal constituent, and we shall also rule these out of the current study.

1.4 The Structure of the Rest of the Thesis

The structure of the rest of this thesis is as follows.

Chapter 2 takes a look at previous research on the functions of cleft constructions in discourse. We review a variety of claims to the effect that cleft constructions serve to mark particular kinds of information, notably FOCUS, GIVEN and NEW, in particular syntactic positions. For most of these claims, it is possible to draw counterexamples from the data, and we demonstrate the problems of a single-layer approach to the description of the function of clefts with these as evidence.

Chapter 3 contains a syntactic analysis of each of the three cleft types in turn, couched in

the terms of Generative Phrase Structure Grammar (cf. Gazdar *et al.* [1985]). We pay particular attention in this chapter to syntactic factors that may serve to differentiate between the three cleft types in terms of their possible discourse functions. In addition, we compare the variety of clefts that are judged in general to be grammatical, and that can be generated by the rules we give in the chapter, to the range of strings actually occurring in the corpus of data.

In chapter 4 we turn to the analysis of the aspect of clefts that is most important in differentiating them from non-cleft constructions: their semantics. We demonstrate that cleft constructions are presupposition-bearing on a fairly conventional description of presupposition, and assess the relative merits of three particular theories of presupposition for describing clefts. The account we adopt is that of van der Sandt [1988 and forthcoming], who suggests that a fruitful way to look at a presupposition is in terms of an analogy with the treatment of pronominal anaphora within the framework of Kamp's [1981] Discourse Representation Theory. Accordingly, we work out the details of how van der Sandt's proposal for presupposition in general might work for clefts in particular. In this chapter, we also touch on the important issues of the referentiality of all or parts of clefts, and examine the claim that cleft heads in particular specify semantic objects that are *UNIQUE* with respect to their ability to satisfy the existentially-quantified variable in the cleft presupposition.

In chapter 5 we move on to an account of some separate but related issues concerning the use of clefts in discourse. We examine first of all the range of accentual patterns available to clefts, further substantiating claims we made in chapter 2. We then look at the range of relationships it is possible for cleft presuppositions to have with the discourse context, following our observations in chapter 4 to the effect that the presupposed information need not always be currently available to the hearer prior to the time of encountering the presupposition. We identify three types of presuppositional relation, and examine the profile of each cleft type in a corpus of 150 examples with respect to the frequency with which particular presuppositional relations appear. On this basis, we are able to establish distinct presuppositional profiles for each type of cleft.

We move on to an analysis of the information structure of the cleft as a whole, using the framework developed for the discussion on presupposition types as a basis. We establish

both general features of cleft information structure, and some further factors by which clefts are differentiated. In the final section of the chapter, we point out the distinct and complementary roles of information structure and cleft presupposition, showing how they combine to determine a range of functions unique to clefts. We also discuss parallels with the model for human sentence processing suggested by Clark and Haviland [1977], indicating the potential of that model for explaining some factors observed in the corpus with respect to the information structure and presuppositional behaviour of the data.

Chapter 6 presents a summary of the conclusions reached in the thesis, and discusses how they might be applied in a decision procedure for choosing whether to cleft, and which cleft to use. Finally, some suggestions are made with respect to further relevant research.

Having outlined the aims and content of the thesis, we can move on to the first part of the research: a review of previous work on the discourse functions of clefts.

Chapter 2

An Overview of Previous Research

In this thesis, we are particularly interested in why people use clefts—that is, with the reasons behind the choice of a cleft construction given a particular context and a particular message to convey. In chapter 1, we explicitly adopted the assumption that cleft constructions belong to a class of syntactic structures whose use bestows particular pragmatic markings on the message being carried, these markings being a powerful determinant of syntactic choice. This chapter examines claims that have been made in the literature regarding the nature of those markings.

An examination of the literature reveals that claims regarding clefts fall into two broad categories. Those in the first category have in common the view that the use of a cleft construction, either of itself or in conjunction with a particular accent pattern, indicates that the speaker or writer considers or intends certain elements within the construction to be interpreted as FOCAL. Unfortunately, although intuitive characterisations of the meaning of the term FOCUS are many and varied, definitions are few and far between. We will assess the various claims regarding the focusing functions of clefts in section 2.1: as we will see, the major thing these accounts have in common is the use of the term FOCUS, rather than a common understanding of what the term means.

The second category of claim, addressed in section 2.2, consists of accounts taking the view that information types other than focus are marked by cleft constructions. The information types invoked include the notions GIVEN, NEW and KNOWN information.

Some of these accounts discuss the function of only parts of the cleft; others go further in trying to account for the function of the construction as a whole. The latter is generally done in one of two ways, the claim being either that clefts are a means of presenting information in a particular serial order, or that they indicate a distinction between two different types of information, usually GIVEN and NEW.

2.1 Accounts Based on Focus-Marking

In this section, we examine claims to the effect that clefts can be explained by some formulation of the notion of FOCUS. Many things are meant by this term: in particular, the term 'focus' can either denote a syntactic or prosodic INDICATOR of information status, or it can be a term predicated of the information itself, independently of how focal status came to be associated with it. In one case, Quirk *et al.*'s [1985] account, sentence ACCENT is itself termed 'focus'. In Geluykens' [1984] account, accent is taken to be the indicator of focal information; Sidner [1979], Reichman [1981] and Taglicht [1984] take syntax to be the indicator of the focal information; and Creider [1979] and Chafe [1976] take the view that accent and syntax together indicate focus.

Taglicht's claim that the cleft head is focus will not be examined here, since he makes no predictions as to what will result from his analysis: his syntactic focus can only be taken as an alternative label for the syntactic constituent concerned. Reichman and Sidner, on the other hand, in their computational accounts of the relationship between focusing and anaphora, offer a testable hypothesis—that is, one that provides both a definition of the notion concerned, and a claim regarding its purpose.

It is not my intention here to discuss the relative merits of different notions of focus *per se*; we will be concerned only with the applicability and usefulness of particular notions for describing the functions of clefts in discourse, including the accuracy of any predictions they make with respect to the surrounding discourse. Throughout the discussion, since there are so many different notions of focus being dealt with, it should be assumed that where the term 'focus' is used, the notion intended is that of the author being discussed at the time.

2.1.1 Chafe

Focus of Contrast

Chafe [1976] distinguishes a class of sentence types he calls CONTRASTIVE SENTENCES; these, he argues, mark a distinction between BACKGROUND KNOWLEDGE and FOCUS OF CONTRAST. Background knowledge is 'that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance' [Chafe 1976:30], and is 'pronounced with lower pitch and weaker stress than new' [Chafe 1976:31]; focus of contrast, on the other hand, is expressed through 'higher pitch and stronger stress' than the other material in the sentence. This focus (there is by obligation one, but may be more, in a contrastive sentence) is described as the ASSERTED ALTERNATIVE to some other element that a hearer may have in mind, where both the alternatives are potential candidates for satisfying the predicate in the background knowledge of the sentence.

For example, in the following short discourse (taken from Chafe [1976:25]), *Sally* and *Ronald* are contrastive alternatives:

- (2-1) a Sally made the hamburgers.
b RONald made the hamburgers.

Clefts as Contrastive Sentences

Chafe makes two related claims for the discourse function of clefts. First, he suggests that they can be used to mark focus of contrast, where the focus of contrast appears on the cleft head. Chafe's own position is a reiteration of Jespersen's:

A cleaving of a sentence by means of *it is* (often followed by a relative pronoun or connective) serves to single out one particular element of the sentence and very often, by directing attention to it and bringing it, as it were, into focus, to mark a contrast.

[Jespersen 1961:147–148]

Chafe's examples of this use of clefts are as follows:

- (2-2) It was RONald who made the hamburgers.

(2-3) The one who made the hamburgers was RONald.

Chafe's second claim is that

...cleft sentences provide a way of moving focus to the the right, so that it does not appear as the initial item in the sentence. There may, then, be a tendency to place not only new information, but also the focus of contrast later in the sentence. Pseudo-cleft sentences, of course, allow this tendency full rein.

[Chafe 1976:37]

In the cases shown in (2-2) and (2-3) above, focus of contrast is, as Chafe describes, further to the right than it might have been in the corresponding declarative, shown in (2-4):

(2-4) RONald made the hamburgers.

Of course, this account, and other accounts based on the 'movement' of particular types of information to particular places in the cleft, presuppose that a corresponding declarative is available for the cleft in question. This is not always the case, however, since what we might take to be the corresponding 'unclefted' constructions are sometimes syntactically unacceptable. It can be very difficult to determine exactly what the unmarked declarative correlate of a particular cleft should be (perhaps this is evidence in itself that there is something suspicious about accounts based on movement), but in many cases all the alternatives that present themselves are equally unacceptable. For example, the declarative corresponding to the reverse *wh*-cleft:

(2-5) That's who it is.

is presumably one of the following unacceptable sentences:

- (2-6) a * It is that.
b * It is that person.

Halvorsen [1978:6] comments on a number of other cases where cleft sentences have no declarative correlate. This is evidence for a model of the choice to cleft that takes into account not only pragmatic considerations but syntactic constraints on possible linguistic realisations as well.

Some Problems

Apart from the point just mentioned, Chafe's account also runs into problems when confronted with other clefts data.

First of all, examples in which there does not appear to be anything corresponding to focus of contrast on the cleft head appear with regularity: (2-7) is one such example, where the cleft head is unaccented, and does not seem to represent the asserted alternative that Chafe's account requires:

- (2-7) a Joe Wright you mean.
b Yes yes.
c I thought it was Joe Wright who'd walked in at FIRST.

As a corollary of this, it is clear that elements that fit the definition of focus of contrast may appear in parts of the cleft other than the head. Such examples are not a problem for Chafe's account provided these foci appear in addition to a focus in the head, since he states that there may be multiple foci of contrast in a single sentence.¹ However, note that, in (2-7), the sole focus of contrast, as defined in prosodic terms, appears to be in the complement.

As a second problem, we might note that Chafe makes no comment about the occurrence or function of accents which do not appear to indicate contrast: these often appear in *it*-cleft complements. (2-8b) is one example where such an accent appears (secondary accent is indicated here by means of italics):

- (2-8) a I've done the shopping.
b Was it your turn? I thought it was MIKE who took the *list*.

The information in the complement of (2-8b) cannot be described as either an asserted alternative or as Chafe's background knowledge. In fact, it is quite simply new but non-contrastive information, which does not figure in Chafe's division of information in clefts.

Third, difficulties also arise with Chafe's claim that clefts serve to move the focus of contrast to the right in an utterance. Even if it is assumed that focus of contrast occurs

¹In fact, it is not clear whether Chafe intends that this claim be taken to extend to clefts.

where he predicts, his claim does not account for clefts with heads that would be the grammatical object of the corresponding declarative. For example, it cannot be claimed that the function of (2-9b) is to postpone the contrast, since the cleft places the focus considerably further to the *left* than would be the case were the declarative (2-9c) to be used instead:

- (2-9) a I've bought Martha's rabbit. You'll have to wrap it up.
b But it was a GUINEA pig I wanted you to buy!
c But I wanted you to buy a GUINEA pig!

Summary

It is clear from the evidence above that Chafe's suggestions are not sufficiently general to account for the kinds of cleft he sets out to explain. The difficulties encountered show first of all that Chafe's account of clefts as 'contrastive sentences' in the sense he intends cannot be supported. In addition, the data we have seen so far are sufficiently heterogeneous in accentual pattern to suggest that a motivation of clefts in terms of the placement of any kind of accentually-marked information, contrastive or otherwise, is unlikely to suffice as an account of clefts in general.

In chapter 5, we show that there are no less than five predominant accentual patterns for clefts, and that a large proportion of the data serve as counterexamples to Chafe's claims.

2.1.2 Creider

Creider's [1979] proposal is based on the argument that many languages clearly differentiate FOCUS from TOPIC. He argues that some syntactic structures are focusing constructions, and that others are topicalizing constructions; *it*-clefts are said to fall into the category of focusing constructions.

Creider's notion of focus is essentially based on sentence accent, although he establishes the notions of topic and focus on the basis of three tests, as follows.

The Assertability Test for Focus

First, a constituent in a given utterance is said to be the focus if it can be construed as being ASSERTED in that utterance [Creider 1979:9]. Although, as Ward [1985:31] has observed, the notion of ‘asserting a constituent’ is somewhat difficult to conceptualise, Creider’s test is that a constituent is being asserted in an utterance, and is therefore the focus, if the utterance provides an answer to a *wh*-question by identifying that constituent as the instantiation of the *wh*-word in the question. Creider’s [1979:4] examples, repeated in (2-10) below, show that a declarative, such as (2-10b), can be focusing in this sense, while a topicalisation, such as (2-10c), cannot:

- (2-10) a What kind of breakfast food can you eat every morning?
b I can eat English muffins every morning.
c # English muffins I can eat every morning.

The *about*-Test for Topic

The second test delineates Creider’s TOPICS as elements that can be introduced in an *about* phrase, and then referred to subsequently. Creider’s [1979:5] example appears in (2-11); it shows that the appropriate topicalisations, such as (2-11b), provide a topic, while (2-11c) and (2-11d) do not:

- (2-11) a What do you find exciting about Griselda?
b Griselda, I hope to meet her husband someday.
c #I hope to meet Griselda’s husband someday.
d # Griselda’s husband, I hope to meet him someday.

The Non-negatability Test for Topic

Creider’s third test for differentiating between his notions of topic and focus is that when the topic is sentence-initial, it cannot occur within the scope of negation, as he demonstrates using the following examples:

- (2-12) a # It is not the case that the play, John saw it yesterday.
b #It is not the case that as for the play, John saw it yesterday.

On the other hand, sentence-initial focus (in the following example, *Raymond*) *can* appear within the scope of negation:

(2-13) It is not the case that Raymond made the hamburgers.

Clefts as Focusing Constructions

Creider claims that, in English, transformations that effect movement to the *left* are TOPICALISING, and those that move constituents to the *right* are FOCUSING.

On Creider's account, however, although *it*-clefts belong to the class of structures involving leftward movement (presumably because the interposing of the cleft complementiser could be seen as causing the cleft head to move to the left), they are an exception to the rule that correlates movement with focusing or topicalising, since they appear to be focusing in spite of belonging to the leftward-movement category.

Creider states that in *it*-clefts the 'stressed constituent' is the focus. It might therefore be assumed that any stressed constituent anywhere in a cleft construction should perform appropriately on the basis of the tests for focus and topic given above. However, Creider goes on to equate 'stressed constituent' with the cleft head in all but what he considers to be a limited number of exceptional cases. He states:

Note that cleft sentences ... are type cases where the stressed constituent is focused. It is impossible to read (i) below without stressing the underlined constituent:

(i) It is John who likes 'An Goirtin Eornan'.

The only situation in which *John* would not be stressed would be one where a correction was being made with respect to *John*, as in (ii) (with focusing stress on the portion in boldface ...):

(ii) It is John who likes 'An Goirtin Eornan'.

It is inherent to metacontexts such as these that normal values be reversed.

[Creider 1979:15]

Problems

As we saw above in relation to the Chafe's explanation of clefting, the heterogeneity of clefts with respect to accentual patterns is clearly greater than Creider's characterisation suggests (we go on in chapter 5, moreover, to show that the accent pattern Chafe claims to be 'normal' for *it*-clefts is in fact the least usual of the patterns that occur). As a result of this, two of Creider's tests for focus and topic do not support his conclusion that *it*-clefts in general are 'focusing', as follows.

In support of his claim that clefts are 'focusing', Creider gives an example where a cleft with the 'normal' accentual pattern does not pass the *about*-test for topic:

- (2-14) A: What is it about John that makes him easy to work with?
B: # It's JOHN that writes all the lines.

However, it is possible to construct examples with the same accentual pattern where the cleft does pass the test for topic:

- (2-15) a What is it about Mike and Sandy that you like?
b It's only SANDY that I like.

It is also possible to construct examples with different accentual patterns which pass the *about*-test. (2-16), for example, has two nuclei:

- (2-16) a You'd better tell me about this guy Pete.
b It's HIM who's CAUSED all this trouble, after all.

Clefts with weakly accented or unaccented heads can pass, too, thereby qualifying these elements as topics. This is demonstrated in (2-17):

- (2-17) a What is it about John that makes him easy to work with?
b It's John who writes all the LINES, and so he can give us a lot of help with interpretation.

The *about*-test shows, therefore, that some cleft heads are apparently topics, on Creider's definition.

A second problem arises with Creider's prediction that clefts with the 'normal' accentual pattern will pass the *wh*-question test for focus. As we noted above, the focus is that

element which can instantiate the *wh*-word in a *wh*-question:

- (2-18) a Who did the washing?
b It was JOHN who did the washing.

Clefts with accentual patterns other than that exhibited by (2-18b), however, fail this test for focusing constructions. This is because clefts with unaccented or weakly accented heads cannot be used as direct answers to *wh*-questions at all: the resulting discourse is incoherent. To demonstrate how such patterns occur in contexts, consider (2-19b), which is an *it*-cleft with an unaccented head and accented complement:

- (2-19) a You should complain to John.
b But it's John that's CAUSING this!

When the cleft in (2-19b) is placed in a question context such as the focus test would require, where the head of the cleft would act as the instantiation of the *wh*-word in the question, the result is unacceptable:

- (2-20) a Who's causing this?
b # It's John that's CAUSING this!

These clefts, therefore, do not 'focus' in the way Chafe suggests.

Ultimately, Creider's assumption that clefts are 'focusing' on his definition results from too narrow a sample of data. His tests for focus seem to capture clefts with heads that carry new information, and the topic tests capture clefts with contrastive, or given heads.² Creider's account founders, like Chafe's, chiefly as a result of the assumption that an accented constituent necessarily appears in the cleft head. For this reason, Creider's tests do not show that clefts in general are focusing constructions, contrary to

²It might at this point be tempting to conjecture that the clefts that pass Creider's test for topic are those which fall into Prince's [1978] class of informative-presupposition (IP) clefts, and those that pass the focus test are stressed-focus (SF) clefts (this distinction is discussed in section 2.2.1). This conclusion would, however, fall foul of Creider's final test for topic and focus, which states that initial focus can fall within the scope of external negation, but that initial topic cannot. This test does not discriminate between *it*-clefts of the IP type, such as (2-21), and those of the SF type, such as (2-22):

- (2-21) The railwaymen saw the incident, admittedly, but it is not the case that it was they who rePORTed it.

- (2-22) Someone must have done it, admittedly, but it is not the case that it was PAULine.

his claims.

2.1.3 Quirk *et al*

There are two strands to Quirk *et al.*'s [1985] argument about clefts. The first concerns the relationship between focus and accentual pattern, and is discussed below. The second relates to the role of the cleft as a syntactic indicator of the boundaries of given and new information in the structure, which we discuss in section 2.2.4 below.

Focus and Intonation

Quirk *et al.* [1985:1353ff] subscribe to a notion of focus akin to that of Halliday [1967, 1973, 1985], where focus is a linguistic notion defined on the basis of prosody.

Human speech is naturally segmented into brief stretches which, in general, correspond to units of information. Each such INFORMATION UNIT is realized as a pitch contour called a TONE UNIT (or TONE GROUP, in Halliday's [1985] terms), so that information units and tone units are co-extensive. Each tone unit contains one INTONATION NUCLEUS, this being the peak of greatest prominence in that unit.³

On this account, FOCUS indicates 'where the new information lies' [Quirk *et al.* 1985:1363] in an information unit, and is signalled by the intonation nucleus. In Halliday's [1985:275] terms, the element of the utterance which has this TONIC PROMINENCE is said to be carrying INFORMATION FOCUS. The new information in question can be as little as a single syllable, or as much as a whole clause.

Focus in Clefts

In the case of cleft constructions, Quirk *et al.* argue that the traditional correspondence between information unit, new information, and nucleus is fundamentally altered, as follows. The claim is that a cleft consists of a single information unit, but that the function of the cleft is to 'divide' the one focus into two, resulting in a single unit with

³For an elaboration of these concepts, see, for example, chapter IV in Couper-Kuhlen [1986].

two foci. Quirk *et al.* therefore account for *wh*-clefts and *it*-clefts as devices for achieving DIVIDED FOCUS [1985:1372]. This division is claimed to fall at the syntactic boundary between the head and the complement of the clause, and Quirk *et al.* predict that a nuclear accent will fall at the end of each constituent, as in (2-23):

(2-23) It was JOHN that had him WORRIED.

In spite of there being two nuclei, both of which on the standard assumption would be associated with new information, on Quirk *et al.*'s account only one of these two nuclei is expected to indicate new information. In such cases (for the cleft is not the only construction which allows divided focus) the nucleus associated with the new information is termed the DOMINANT NUCLEUS, and the constituent bearing that nucleus will be the dominant constituent. In the case of the *it*-cleft, the context determines which of the two nuclei appears on new information, and so either can be the dominant one [Quirk *et al.* 1985:1384]. In the case of the *wh*-cleft, on the other hand, the dominant focus is assumed always to occur on the head.

Discussion

Quirk *et al.*'s account is interesting in that, in its claims for accentual structure, it comes a great deal closer to accounting for the range of clefts data than the other accounts expressed in terms of focusing that we examine here. In chapter 5, we will see that clefts containing two nuclei are very common, particularly in the case of the *wh*-cleft. Quirk *et al.* also avoid the common pitfall of taking an accent-based view of focus and then supposing that this will correlate reliably with syntactic structure, as has been the case in other accounts examined in this section. Their account, however, runs into several problems, as follows.

First, the analysis Quirk *et al.* suggest implies a revision of the established notions of information unit and tone group of the most fundamental kind. In the standard accounts (as found, for example, in Crystal [1969]), it is axiomatic that a single tone group has a single nucleus, and the suggestion that two nuclei are possible suggests that the unit Quirk *et al.* postulate is something very different from the traditional notion. However, they provide no definition that would make clear what their new notion consists of.

Moreover, Quirk *et al.* fail to argue convincingly for the controversial claim that nuclei can correlate with non-new information. In addition, the analysis of clefts as consisting of a single information unit, and therefore as a single tone group, is hard to reconcile with the phonological evidence: for example, Geluykens [1984] reports *wh*-clefts consisting of up to 17 conventional tone groups, and much of the data on tone groups presented in chapter 5 of this thesis suggests that ‘single group’ clefts are in a minority.

Second, even if Quirk *et al.* were able to show that their new analysis is justified, their account again falls short in terms of coverage of the data, since the explanation of clefts requires that all clefts have a double nucleus. This means first of all that clefts with more than two nuclei cannot be accounted for. In addition, clefts with fewer than two nuclei are also ruled out, which, interestingly, disallows the kinds of *it*-cleft that many researchers (including some discussed in this section, namely Chafe [1976] and Creider [1979]) consider to be the only common types. As we noted above in relation to Creider’s claim, single nucleus accent patterns are not the most normal, but they certainly occur. Examples of each appear in (2-24):

- (2-24) a It was the WIND that knocked off her hat.
 b What knocked off her hat was the WIND.

Although Quirk *et al.*’s description of the function of clefts does come closer to coverage of the data than the others discussed so far in this section, the evidence above shows that their account still falls short of a complete description. In addition, their view implies unprincipled revisions of basic notions in phonology and information structure that require a great deal more support if they are to be convincing.

2.1.4 Two Theories of Focus in Computational Linguistics

In this section, we consider two computational approaches to the tracking of focus in discourse that have something to say about clefts: Sidner’s [1979] theory of focusing for anaphora resolution, and Reichman’s [1981, 1985] model for the generation and interpretation of conversational discourse. The two theories are unique among the research involving focus being dealt with in this chapter in that, while the other accounts provide definitions of focus, no predictions are made with respect to what focus as defined in each case actually *does* in discourse—that is, what effects, linguistic or otherwise, are

meant to result from an element being in focus. Sidner's and Reichman's theories, on the other hand, both contain notions of focus that were evolved as a means of capturing the behaviour of discourse phenomena, and can be tested by assessment of how well they account for that behaviour.

Although the two theories differ in how focus is defined and exploited, in both cases focus is a status that is predicated of information in a model of the ongoing discourse, and is related directly to the use of anaphoric referring expressions, the hypothesis being that focal elements are the most likely antecedents for subsequent anaphoric expressions.

Reichman

In Reichman's [1981, 1985] model, a discourse is represented by a collection of **CONTEXT SPACES**, each of which corresponds to a particular segment of the discourse; that context space which corresponds to the current part of the discourse is referred to as the **ACTIVE** space. Each context space contains, amongst other things, the entities mentioned in the part of the discourse corresponding to that context space; and each entity has a corresponding **FOCUS LEVEL** which indicates its salience within that context space. For the speaker, the focus level assigned to an entity determines the form of referring phrases to be generated; for the hearer, the focus level assists in the interpretation of anaphoric expressions.

Focus levels are established by means of a set of 14 **FOCUS RULES**. These predict the changes that will take place in the focus registers of the currently active space either after generating or interpreting an utterance with particular characteristics. For example, a semantic rule exists that stipulates the assignation of a **HIGH** focus level to the agent of an action. The syntactic rules concern the use of particular kinds of constituent or sentence construction, and include the following rule for clefts [Reichman 1985:75]:

F3 The subject of a pseudo-cleft, cleft, or topicalised clause is assigned a high focus assignment.

In Reichman's model, only high focus elements in the currently active context space may be referenced by a pronoun [Reichman 1981:119]. Her model therefore predicts

that when an *it*-cleft or reverse *wh*-cleft occurs, its head is a likely source of antecedents for subsequently-occurring pronominal anaphora; when a *wh*-cleft occurs, the likely source is the *wh*-clause.

Sidner

Before examining the accuracy of Reichman's prediction, we will look briefly at a similar assumption made by Sidner [1979], and implemented in her program for the resolution of anaphora. Sidner claims that clefts are in a class of syntactic constructions 'which make the recognition of focus easy, since these sentence types are claimed to have the purpose of singling out one element from others' [Sidner 1979:60]. Sidner's notion of focus is similar to Reichman's, in that it is a dynamically-updated assignment of salience to entities that responds to linguistic aspects of the input. The model as a whole differs in that Sidner takes that input to indicate **POTENTIAL** foci—that is, likely antecedents for pronouns—but does not decide on actual focus until this has been indicated by the resolution of an anaphor. Sidner's potential foci are therefore arranged in a list according to rules based on her theory of the likelihood of their being successful candidates for actual focus. The heads of clefts of all three syntactic types are claimed to have the function of placing the element referred to by the cleft head at the top of the list, with the result that this element—the **EXPECTED FOCUS** in Sidner's terms—is the first candidate to consider when attempting to resolve a subsequent anaphor.

Discussion

It is not our purpose here to criticise the status of Reichman's and Sidner's notions of focus either in terms of their 'correctness' or in terms of the degree to which they approximate what is intuitively meant by 'focus'. Both are formally defined notions with an explicit function in a computational theory, and to this extent other uses of the term, and intuitions about what it should mean, are irrelevant.

What is of interest is the value of the inclusion of clefts as focusing constructions in theories of anaphora resolution. In each case, we are interested in exploring whether the rule for clefts can really enhance the efficiency of the algorithm in question. Sidner's

Cleft Type	Antecedent		Total
	Head	Non-head	
<i>it</i> -cleft	10 (43%)	13 (57%)	23
<i>wh</i> -cleft	32 (78%)	9 (22%)	41
reverse <i>wh</i> -cleft	3 (11%)	25 (89%)	28
Total	45 (49%)	47 (51%)	92

Figure 2.1: Clefts as providers of antecedents

claim can be tested in the form of the hypothesis that the heads of cleft constructions provide antecedents for subsequent pronominal anaphors at a rate that is better than other parts of the cleft construction; Reichman's claim can be tested on the same hypothesis for *it*-clefts and reverse *wh*-clefts, and on the hypothesis that *wh*-cleft *wh*-clauses provide antecedents more frequently than *wh*-cleft heads.

To test these hypotheses, I took a sample of 150 sentences, 50 each of *it*-clefts, *wh*-clefts, and reverse *wh*-clefts, from my corpus of naturally-occurring spoken data. 92 of these had one or more subsequent pronominal anaphors within three sentences of the cleft itself.

The antecedents of these anaphors were categorised as being either HEAD—that is, occurring in the head of the cleft—or NON-HEAD—occurring either in the cleft complement, or being composed of the entire content of the sentence, as sometimes occurs with the anaphoric elements *this* and *that*. The results of the analysis appear in figure 2.1.

The distribution of pronoun resolutions shown in figure 2.1 is significant to .001 by the Chi-Square test. Taking Sidner's hypothesis first, the table shows that for two out of three of the cleft types—*it*-clefts and reverse *wh*-clefts—there is unlikely to be any advantage in taking the cleft head as a favoured antecedent for subsequently-occurring pronouns; in the case of the reverse *wh*-cleft, it is likely to be a distinct disadvantage. In Reichman's case, again, *it*-cleft heads offer a roughly equal chance of providing an antecedent and not doing so; and reverse *wh*-cleft heads provide a much worse likelihood. *Wh*-cleft *wh*-clauses also provide a much smaller number of antecedents than *wh*-cleft heads. Reichman's theory therefore fares slightly worse than Sidner's, although neither fares particularly well.

While it is not in question that Sidner's and Reichman's algorithms do resolve anaphors, on the basis of the evidence above we would question whether the view of pragmatically-marked syntactic constructions that is built into them is actually of assistance in the task.

Of course, the view of the function of marked syntactic structures adopted in Sidner's and Reichman's work is necessarily drawn from the linguistics literature, and is likely therefore to give currency to views originating there, some of which are criticised in this chapter. Undoubtedly, though, on the basis of work such as Sidner's and Reichman's, the view of clefts as focusing structures as the term is meant here has become widespread in the natural language processing community and its literature (see, for example, Carter [1987:104]). The data above suggest that a review of this assumption is in order.

2.1.5 Geluykens

Focus as New or Contrastive Information

Geluykens' [1984] notion of focus is similar to Halliday's definition of new information. Geluykens describes focus as 'information which is New and/or Contrastive' [Geluykens 1984:27]; New information is

...information which the speaker does not assume to be in the hearer's consciousness, and which he also cannot assume to be recoverable by the hearer from the discourse context.

Geluykens [1984:23] explains the functions of *it*-clefts, *wh*-clefts and reverse *wh*-clefts in terms of the view that all three are special devices for the marking of focal material as he describes. Focus is not only marked by syntactic constructions, however:

Focus will be expressed through Prosody, Linear Order, or Special Devices, or through any combination of these three phenomena.

[Geluykens 1984:28]

Geluykens' study is intended to determine the relative importance of cleft constructions, prosody, and linear order in marking his notion of focus. He proceeds first of all by

dividing his corpus of clefts, which consists of examples of all three syntactic types, into subcategories depending on where the focal material occurs. Since it is clear that this focal material is not always present in the head of the cleft, Geluykens examines the data with the aim of explaining its placement elsewhere in the sentence. He does this using a set of seven separate hypotheses, his purpose being to find out how successful each of them is in accounting for the data. The seven hypotheses are as follows:

1. Given information should appear before New information.
2. Focus (as defined above) should appear near the end of the sentence.
3. There should be more New information than Given in any sentence
4. Information may need to be moved into, or out of, 'Topic Position'.
5. svo order should be preserved.
6. Canonical order may need to be disrupted for special effects.
7. Contrast may be required, or may need to be avoided.

Geluykens observes that these principles are able to account for the position of focus in the data to varying degrees, as summarised below.⁴

The Given-before-New Hypothesis

The Given before New principle is as follows:

Given information tends to precede New information in the linear order of the sentence.

[Geluykens 1984:22]

We noted Geluykens' definition of New information above; Given, for Geluykens, is

...information which may be assumed, by the co-operative speaker, to be appropriately in the hearer's consciousness and/or which may be assumed to

⁴It should be obvious, of course, that these principles are not mutually exclusive: some subsume others.

be recoverable by the co-operative hearer from the discourse context (either textual or situational), either more or less directly, or via inferential linking.

[Geluykens 1984:20]

Under these definitions, the claim that Given information generally precedes New information accounts for many *wh*-clefts and most reverse *wh*-clefts; but Geluykens notes that it cannot account for *it*-clefts that present new information first, such as (2-25):

(2-25) A: Who did this?

B: It was JOHN who did it.

Of this New-before-Given ordering, Geluykens observes that there are quite severe restrictions on the *length* of the New information. He suggests that the shortness of the head in cases such as (2-25) is due to the reluctance of the speaker to postpone the Given information for too long.

The End-Focus Constraint

Since Geluykens takes the view that Focus and New information frequently coincide [1984:27], this constraint is almost the corollary of the Given-before-New hypothesis: it is simply that the Focus tends to come near the end of the sentence.

The End-focus hypothesis is seen to be supported by the *wh*-cleft, because of the position of the head, commonly accented, at the end of the sentence, as in (2-26a).

(2-26) a What I found in the barn was THIS.

b I found THIS in the barn.

In addition, 40% of the *wh*-clefts in Geluykens' corpus have direct objects as head, which the *wh*-cleft is able to postpone to the end of the sentence without this effect being diluted by succeeding indirect objects, complements, etc. as it might be in a declarative such as (2-26b).

More New Information than Given

The More-New-than-Given hypothesis states, predictably, that an utterance will contain more New than Given information. Geluykens observes:

If the Focus is indeed the most important part of the message, we might reasonably expect that the Focus consists of more words than the Given information.

[Geluykens 1984:35]

Geluykens establishes this on the basis of counting the words that are deemed to carry each kind of information. He does not find, however, that in clefts on average the principle is supported, and so the principle is ruled out by Geluykens.

The Preservation of SVO Order

In the description of several of his cleft types, Geluykens maintains that the type of cleft chosen depends to some extent on whether or not the speaker wishes to preserve the svo order of the sentence, while at the same time achieving some sort of prominence on the desired element by means of some other principle such as end-focus.

The *it*-cleft has a high proportion of subject heads, and so Geluykens concludes that in general it supports the svo ordering hypothesis. *Wh*-clefts appear to support the svo hypothesis, since many of the heads are objects; however, in the case of subject heads (14.17% of Geluykens' corpus) and some adverbials (6.67%) the svo hypothesis fails. Finally, in the case of the reverse *wh*-cleft, the svo hypothesis is not in general supported, as many of the heads (sentence-initial in these sentences) are grammatical objects.

Moving Into or Out of Topic Position

The term TOPIC is used here to indicate

...what the message is about, the pragmatic point of departure of the message ... Topic is not equated with 'Given', nor with 'first element of the clause', although these factors often coincide.

[Geluykens 1984:21]

In Geluykens' descriptions of particular types of cleft, however, he refers to elements being topicalised or de-topicalised by movement to the left and right respectively. This suggests that, for Geluykens, the relevant definition of topic is 'first element of the clause'.

Geluykens concludes that, for the *it*-cleft, detopicalising is an important consideration, since subject heads are moved further to the right than would be the case in the corresponding declarative. Reverse *wh*-clefts, on the other hand, support topicalising, since the head element is in leftmost position in the sentence. For some types of reverse *wh*-clefts, the principle would allow topic and focus (on Geluykens' terms) to coincide: Geluykens suggests this might provide a reason for the speaker to choose a reverse *wh*-cleft rather than the comparable *it*-cleft, as in the latter case the topic and focus are on distinct elements.

Movement from Canonical Position

The point here is simply that a 'marked' reading will occur if an element is moved out of its canonical position in the sentence—for instance, a rightward movement of the cleft head in the case of subject-headed *it*-clefts. Such markedness would presumably be used to achieve prominence for a particular element or for the message as a whole. Geluykens concludes that this principle does not fare sufficiently well in the explanation of the data to feature in the final analysis, however, and in any case it seems that at least part of its function is subsumed by the 'End Focus' and 'Topic Position' principles described above.

Avoidance of Contrast

Geluykens [1984:37] puts forward the hypothesis that in some cases a nuclear accent on a particular element in a declarative (for example, on a subject) would result in a contrastive reading for that element where none was required. A cleft might therefore be used to mark prominence without contrast. However, Geluykens notes that the heads of at least one kind of *it*-cleft are contrastive anyway; on this basis, he discounts this particular hypothesis.

Discussion

Of the seven hypotheses Geluykens suggests as explanations for the arrangement of information in clefts, he discounts three himself. Those remaining are the ones concerned with the linear ordering of elements in the sentence: the Given-before-New and End-Focus hypotheses, and the explanations concerning svo ordering and topicalising. Geluykens notes that the Given-before-New hypothesis is fairly well supported, although as we saw above some *it*-clefts have New-before-Given structure. This is similarly the case with the End-Focus hypothesis, with the same set of counterexamples.

In addition, neither principle explains *wh*-cleft cases with 'focal' heads that are indirect objects, prepositional phrases, or intransitive verb heads, since these elements would be at the end of the sentence in the corresponding declarative in any case: the principles would suggest that clefting is unnecessary in such cases.

For the svo hypothesis, Geluykens seems to overestimate the support. First, he claims that the high proportion of subject-headed *it*-clefts support the hypothesis. However, *it*-clefts with subject heads preserve svo ordering precisely as much as canonical declaratives, and so the fact that a cleft is used is not explained. In addition, he fails to observe that *it*-clefts with other head types, which make up 47.1% of his *it*-cleft corpus, feature heads that on the model he suggests would have been fronted from a post-verbal position, including adverbials, direct objects, and prepositional complements. These systematically go against the svo hypothesis.

A similar problem arises with the Topicalising hypothesis: again, the conclusion does not

apply to nearly half the data, whose head elements would be described on Geluykens' terms as having been moved not to the right, as Geluykens' topicalising argument requires, but to the left.

It is not clear which of the remaining four hypotheses is the strongest, but it is obvious that none individually constitutes an explanation of clefting. From the discussion Geluykens presents, which it should be recalled is intended to investigate how cleft constructions, as special devices for marking focus, achieve the focus marking function, it is not clear that any picture of the cleft as a *syntactic* focusing device emerges at all. In fact, Geluykens eventually places most emphasis on the function of *accent* as the principal means of focus marking. He states

We even think that prosody might be the most important focus marker in spoken English.

[Geluykens 1984:25]

and later

In more than 99% of our corpus, Focus is marked by tonality and tonicity ...tonicity is a powerful focus marker, more powerful than linear order.

[Geluykens 1984:69]

Prosody, therefore—the accentual marking of focus—emerges as a stronger marking of Geluykens' notion of focus than any of the linear order principles he tests, including the four remaining ones mentioned above. However, Geluykens presents no argument upon which to base the view that the prosody of clefts is any different from that of declaratives. This leaves entirely unsupported the claim that *clefts* are focusing constructions, since Geluykens does not demonstrate that there is anything special about the syntax of clefts that justifies his enduring view that clefts are special devices for indicating his notion of focus.

Although Geluykens' account fails to demonstrate the 'focusing' function of clefts, his work is useful in suggesting the success rate of various principles in explaining the data. In particular, what is interesting about Geluykens' work is that it does show that several

different principles may combine in accounting for the function of a group of clefts. Which principles operate in any given instance, however, may depend on context, and on what the speaker is trying to achieve. On this assumption, a more plausible account of the function of clefts in discourse would be one in which contextual factors could be formulated to explain which principles would be relevant in which contexts.

It is not necessarily the case that the principles Geluykens has chosen will be all and only the ones that are required: it is likely that the three he discards himself, for example, will not be of use; and he fails to consider the presuppositional nature of clefts, which must be fundamental in the choice of whether or not to use a cleft.

2.1.6 Clefts and Focus-Marking: Conclusion

It is clear that none of the accounts of clefts as ‘focusing constructions’ that we have examined are satisfactory, all of them failing to demonstrate that cleft constructions are ‘focusing’ on any of the definitions of focus supplied. In the majority of cases, the claim has been that the focus of the cleft is the cleft head, or the information borne by that constituent. However, we saw above that this part of the cleft does not show any uniformity of function sufficient to satisfy any of the definitions of focus that have been examined in this section. In particular:

- The head element does not always carry sentence accent.
- The head element does not always carry new or contrastive information.
- The head element does not provide antecedents for subsequent pronominal anaphora at a rate significantly better than elements occurring elsewhere in the cleft.

It is interesting to look into the reasons for the failure of the accounts we have discussed: these are surprisingly general. First of all, with the single exception of Geluykens’ account, the research discussed fails to take sufficient account of the range of data involved, particularly with respect to the variety of information structures and accentual patterns clefts display. The temptation to assume that there is a single ‘normal’, or even possible, accentual pattern for clefts undoubtedly begins in the transformational tradition, where data is drawn from decontextualised examples. The accentual pattern

assumed to be unique for clefts is the default one assumed on uttering examples out of context; such examples, however, are not fit data for conclusions regarding the functions of clefts in discourse.

As a result of this underestimation of the data, Quirk *et al.*'s account, which looks exclusively at accent, fails to predict the range of accentual patterns that would be required for a general account. Similarly, Creider and Chafe, supposing an accentual peak to coincide with the head of the cleft—a *syntactic* position—could not deal with cases where the accentual pattern differed.

The view that the focal element in a cleft is the syntactic head, shared also by Sidner and in part by Reichman, appears to be based upon an influential paper by Chomsky [1970] in which he makes a distinction between FOCUS and PRESUPPOSITION in relation to the cleft. The examples Chomsky gives in the main text show the predictable analysis: the cleft head is Chomsky's focus, while the complement is his presupposition. However, a close reading reveals that Chomsky's analysis is based on sentence accent [1970:70], and not on syntax: it is simply that in the examples he chooses to use, what he calls the INTONATION CENTER occurs on the cleft head. He does suggest, in a footnote, that other accentual patterns are available, which (although he does not make this explicit) would require a different analysis of the sentences into focused and presupposing parts.

Of the accounts we have examined, the most successful at dealing with the range of data was that of Quirk *et al.* [1985], which postulated a relationship between a notion of focus and sentence accent, but did not then assume that this two-way correlation would extend to the popular correspondence with syntactic structure.

We conclude, then, that in spite of the popular assumption that clefts are special focus-marking constructions, no account we have examined here has been able to substantiate the claim.

2.2 Accounts Based on the Marking of Other Information Types

In this section, we address accounts of the cleft's function in discourse that are based on the view that clefts mark kinds of information other than focus. The claims made refer to the marking of GIVEN, NEW and KNOWN information, and the view that clefts divide given from new information or enable the presentation of material in a particular serial order.

2.2.1 Prince

Prince [1978] makes a distinction between *it*-clefts and *wh*-clefts in discourse on the basis of the kinds of information each carries in its complement, arguing that this factor plays a major role in determining their distinct functions in discourse. She claims that *wh*-cleft *wh*-clauses carry GIVEN information, characterised as follows:

Information which the co-operative speaker may assume is appropriately in the hearer's consciousness.

[Prince 1978:903]

Information in the complement of *it*-clefts, on the other hand, is claimed to be

not marked as assumed to be in the hearer's mind ...often, though not always, known from the context ...information which the speaker represents as being factual and as already known to certain persons (often not including the hearer).

[Prince 1978:896]

Below, we examine each of these claims in detail.

Wh-clefts

Prince suggests that, although the content of the *wh*-clause in a *wh*-cleft need not have been explicitly included in the previous discourse, if it has not then it is necessary

for the hearer to be able to construct an inferential BRIDGE (cf. Haviland and Clark [1974]) between the information contained in the clause and the preceding discourse. For such bridges to be constructed, Prince claims that it is sufficient that the information is 'appropriate to the situation' [Prince 1978:899]. Prince's characterisation of GIVEN therefore consists of an adaptation of the following formulation due to Chafe, itself based on that of Halliday [1967]:

Given material [is that] which the speaker assumes is already in the addressee's consciousness.

[Chafe 1974:111-2]

Thus, Prince has added to her definition of given (which we saw earlier) the word *appropriately* in order to accommodate the observation that information in the *wh*-clause need not be currently in the hearer's consciousness; it would simply be appropriate to the discourse if it were.

The data on *wh*-clefts Prince presents shows this adaptation to be necessary, since, as we noted above, it cannot be assumed that the content of the *wh*-clause has been made explicit in previous discourse. For example, the following *wh*-cleft conveys what Prince characterises as CONTRAST [Prince 1978:890], where the contrastive relation holds between issues described as *dynamic* and those that are *constant*:

(2-27) Our position is a dynamic one. It will be more and more refined as conditions change during the course of the struggle. What is constant is our commitment to a revolutionary emancipation of Ethiopia.

The contrasting information in the *wh*-clause cannot be assumed to be already in the hearer's consciousness, and so does not conform to Chafe's notion of Given. Prince herself notes that her formulation of GIVEN is too weak to predict which information would be acceptable in *wh*-clauses and which would not, but goes into some detail in characterising kinds of relationship, like that of CONTRAST, that represent major classes of link—and therefore major kinds of bridge—that users of *wh*-clefts seem to be expecting their hearers and readers to be able to make. We will return in chapter 5 to the kinds of coherence link that the *wh*-cleft regularly exploits.

It-clefts

Prince's claim that *it*-clefts carry KNOWN rather than GIVEN information rests on important observations she makes about the nature of the data. Prince divides *it*-clefts into two types, which she differentiates mainly by their information structure but also by their prosody and by certain aspects of their syntax. These are termed the STRESSED-FOCUS (SF) cleft and the INFORMATIVE-PRESUPPOSITION (IP) cleft.⁵

The STRESSED-FOCUS *it*-cleft has a nucleus on the head element (focus, in Prince's terms), indicated in the following example by small capitals:⁶

(2-28) It's HERE I look like Mina Davis.

It should be noted that the SF category subsumes both *it*-clefts with unaccented complements and those with subsidiary accents; examples of each type appear below, respectively:

(2-29) a Who ate the beans?

b It was JOHN who ate the beans.

(2-30) a How come you're still hungry after all that food?

b This is the first thing I've had all day. It was JOHN who ate the *beans*.

The INFORMATIVE-PRESUPPOSITION cleft has a quite different accentual pattern and information structure. Prince's example [1978:901] is as follows:

(2-31) It was 10 years ago this month that young Irwin Vamplew was bopped on the head by a nightstick while smashing windows in Berkeley in order to end the war in Vietnam.

In this type of cleft, a nucleus appears in the complement, accompanying the bulk of the new information in the sentence. As for the discourse effect of this kind of *it*-cleft, Prince claims:

With these sentences, not only is the hearer not expected to be thinking

⁵Perhaps based on a misconstrual of Chomsky's [1970] conclusions in the manner described in section 2.1 above, Prince uses the terms 'focus' and 'presupposition' as *syntactic* labels. Although this is somewhat misleading, we will preserve the terms 'stressed focus' and 'informative presupposition' during the course of the discussion for the sake of simplicity.

⁶This example is from Prince [1978:899].

about the information in the *that*-clause, but s/he is not even expected to *know* it. In fact, the whole point of these sentences is to *inform* the hearer of that very information.

[Prince 1978:898]

Prince concludes that the function of this type of cleft is to mark a particular piece of information as a KNOWN FACT [Prince 1978:899], leading to the frequent use of the device in situations where the speaker does not want to claim responsibility for the fact conveyed by the cleft, implying that it is already well-known to persons other than the hearer. The status KNOWN is

a choice on the part of the speaker of a particular validity-level s/he wishes to ascribe to the utterance.

[Prince 1978:903]

The distinction between GIVEN and KNOWN is therefore intended at least in part to differentiate the functions of *it*-clefts and *wh*-clefts. There are several points to be made in relation to this, however, which show that the distinction is not so clear.

Some Problems

First of all, some of the types of 'given' Prince describes as characteristic of the *wh*-cleft, and therefore types of given information, can appear quite happily in the complements of *it*-clefts: in fact, in some contexts the two are interchangeable.⁷ The category of given that Prince [1978:889] describes as EXPLICIT INFORMATION is one example. The *wh*-cleft below qualifies as explicit information on Prince's terms: the *it*-cleft is an adaptation with identical content in the complement:

- (2-32) a When they say that focus is important, what they are really talking about is accent.
- b When they say that focus is important, it is accent that they are really talking about.

⁷It will be recalled from the introduction that *it*-clefts and *wh*-clefts have been claimed to be interchangeable in general (see, for example, Akmajian [1970]). In cases where other contextual factors cannot decide between them this is indeed the case. Since many of the factors deciding between cleft types are contextual rather than internal to the clefts themselves (as will be demonstrated later in this thesis), the prevalence of claims to their interchangeability can be at least partly explained by the widespread use of de-contextualised examples, where, obviously, such factors are obscured.

Similarly, *wh*-clefts can carry information implicit in the discourse, as shown in Prince's example repeated here as (2-33a). However, (2-33b) shows that *it*-clefts can also perform this function:

- (2-33) a When talk goes on of 'the national purpose', 'the national life', 'the mainstream of the nation', what is being advanced is the erection of these bourgeois dictatorships.
- b When talk goes on of 'the national purpose', 'the national life', 'the mainstream of the nation', it is the erection of these bourgeois dictatorships that is being advanced.

We would suggest that there is in fact an overlap in the kinds of information that *wh*-clefts and *it*-clefts can carry as complement, which casts doubt on the hard and fast distinction between GIVEN and KNOWN. In fact, if we look more closely at the category KNOWN, the heterogeneity of this information type makes it highly likely that it *subsumes* the content of the category GIVEN—simply because KNOWN spans all the information types there are.

The kinds of information appearing in the KNOWN category include, for example, both information completely unknown to hearers and readers, such as that appearing in (2-31), and information already shared and salient, such as that in (2-29). There are also shades in between, in the shape of information that is available to hearers by inference, such as that in (2-27) and (2-30). It seems, therefore, that *it*-clefts are capable of carrying a range of information types spanning a continuum from 'shared and highly salient' to 'known only to speaker'. The interchangeability of the examples in (2-32) and (2-33), however, suggests that part of this continuum at least is shared by the *wh*-cleft.⁸ It does not confirm Prince's view that the two information types GIVEN and KNOWN are qualitatively different.

Instead, it would appear that the KNOWN label for *it*-cleft complements represents an attempt on Prince's part to generalise across the wide range of information types that she observed appearing in *it*-cleft complements. However, this does not appear to be a particularly valid or useful generalisation. There are two points to note in this respect.

⁸Note that I do not claim that, because *wh*-clefts inhabit a part of the information continuum contained within that of the *it*-cleft, an *it*-cleft can be used in any situation where a *wh*-cleft is appropriate. This is of course not the case, since there are many other factors that determine appropriateness, as we demonstrate throughout this thesis.

In fact, the ‘known fact’ effect Prince alludes to, appearing especially in the case of *IP* clefts, is a result of presenting new information as a presupposition in order that it be ACCOMMODATED by the hearer or reader, a phenomenon well known in the semantics literature (cf. Karttunen [1974], Lewis [1979]). We discuss this in detail in chapter 4. We can re-interpret the ‘known fact’ effect as a side-effect of conveying one particular kind of information in a cleft complement, rather than as the motivation behind *it*-clefts as a whole. The interaction between information type and presupposition will be discussed fully in chapter 5.

In addition, the use of the term KNOWN to describe *it*-clefts in which the presupposed information is shared and salient in the discourse, such as is arguably the case in (2-34), serves no apparent purpose:

- (2-34) a Who ate the beans?
b It was JOHN who ate the beans.

Such information is, as we have seen, no different from kinds of information that can appear in *wh*-clefts, and, in addition, it is not intuitively apparent that (2-34) represents a speaker’s attempt to avoid responsibility for the fact conveyed, as Prince claims is characteristic of KNOWN information.

2.2.2 Lapolla

Building Discourse Models

Lapolla [1986] gives an account of the function of *wh*-clefts in discourse in terms of an analogy with another pragmatically-marked construction, namely INVERSION. An example of inversion appears in (2-35):

- (2-35) Outside stood a little angel.

Lapolla is principally interested in the function of the part of the construction that is first encountered by the hearer in each case: the inverted element of inversions, such as *outside* in (2-35), and the *wh*-clause of the *wh*-cleft. For inversions, Lapolla claims that the function of the inverted element is to draw the hearer’s attention to, or to instruct

the hearer to create, a 'location' in their model of the discourse to which the incoming new information can be related [Lapolla 1986:11]. He states:

The claim of this theory is not that the inverted element is old information, though in most cases it is, but that it is the important element with respect to embedding of information. When it is new information it sets up a context to which new information can be related, including the information in its own predicate. When it is old information, it serves to find the correct context in which to embed the information in the predicate.

[Lapolla 1986:10]

When the information used to set up a context is new, this constitutes, on Lapolla's terms, the construction of a new discourse space in the model of the discourse. The *wh*-cleft, he claims, serves a similar, but more restricted function than inversion in this respect, as follows:

the information contained in the *wh*-clause of *wh*-clefts can be equated with old information.⁹ ... It seems that the main difference between inversion and *wh*-clefting is that inversion can create a completely new discourse space and *wh*-clefting cannot.

[Lapolla 1986:13]

He concludes that

one need not propose any new mechanisms for a linguistic, psychological or computational theory to handle clefting but only adopt, or restrict, those proposed for inversion.

[Lapolla 1986:14]

Some Problems

It should be noted first of all that the treatment of *wh*-clefts in general suggested by Lapolla is unlikely to suffice, for two particular reasons.

⁹He claims that this conclusion follows Prince [1978]. In fact, as will be recalled from our discussion of her research, this is rather a rough approximation of her actual claims: Prince equates the information in the *wh*-cleft *wh*-clause with a slight emendation to Chafe's Given, her phrase being 'information which the co-operative speaker may assume is appropriately in the hearer's consciousness'.

First, Lapolla's claim that *wh*-clefts cannot introduce a new discourse space is hard to assess, since he does not define what a discourse space is, or what having a new or an old one might predict in the way of linguistic behaviour. We do know, however, that the *wh*-cleft can be used to present information that is not currently available to the hearer, and can be associated only by inference—not quite the classical use of the term 'old' information. In addition, it is well known that the *wh*-cleft can be used discourse-initially, as in (2-36):

(2-36) What I'm going to talk about today is inversions in discourse.

A discourse space would have to be already open in such cases for Lapolla's claim to be true for the *wh*-cleft. We do not know, however, whether this is possible or not.

Second, his account reveals a serious underestimation of the distinctions between *wh*-clefts and inversion. In particular, simple differences in the syntactic categories of what can be fronted in inversions, and used as head in the *wh*-cleft, indicate that the two are unlikely to be equivalent. In inversions, PPs, locative and temporal adverbials, VPs and prepositional complements can be fronted, although not simple subjects and objects. In the *wh*-cleft, as we will see in chapter 3, simple objects are among the most common, with non-finite VPs, and subjects next, adverbials and PPs being comparatively rare. This means that not all inversions can be replaced by *wh*-clefts, as is shown in the following example:

- (2-37) a In a large forest stood a house.
b *Where a house stood was in a large forest.

The true test of Lapolla's claim as to the equivalence of the two constructions, however, is to replace the *wh*-cleft with an inversion: if the *wh*-cleft performs a subset of the functions of the inversion, the result should be acceptable. We can see straight away that even in cases where both the head (in the case of the cleft) and the fronted constituent (in the case of the inversion both feature) are syntactically permissible, unacceptability often results:

- (2-38) a What Mike is doing is combing his hair.
b ?Combing his hair is Mike.

This type of inversion might succeed when Mike was being *identified* as the person comb-

ing his hair; it appears to be an example of the type of inversion Green [1980] identifies with the speech of sportscasters (such as *Bringing the ball up is Marty Mestemacher*). It is odd here because hair-combing is not a very noticeable activity by which to pick out somebody, and it is hard to imagine a situation where it might be where sportscaster-type locutions would also be appropriate (in the changing-room after the game, perhaps?). But the serious point is that, in the cleft, the same proposition is being used to very different effect: to point out what Mike (a known person) is doing, rather than to identify him. We therefore conclude that the situations in which both the inversion and the *wh*-cleft would be acceptable are rather more restricted than Lapolla supposes, suggesting a greater distinction in their discourse functions than he suggests.

Even though it may be the case that for Lapolla's purposes his mechanism produces the types of *wh*-cleft he wants (although he does not demonstrate this) it cannot be advanced as a general claim that a theory that can handle inversions will necessarily cope with *wh*-clefts satisfactorily.

However, some of Lapolla's intuitions regarding the use of the *wh*-cleft *wh*-clause as 'the important element for the embedding of information' in some model of the discourse are very much in the spirit of Clark and Haviland's [1974] and Clark and Clark's [1977] proposals on the integration of information into discourse models, which we suggest in chapter 5 is itself a plausible model for the processing of clefts.

2.2.3 Werth

Werth [1984] explains clefts in terms of a view of the information structure of sentences due to research in the Prague School tradition, namely the requirement to place the TOPIC of the sentence before the COMMENT, a similar hypothesis to the Given-before-New principle tested by Geluykens. Informally, the topic is the information that is used to connect to what is already shared between the speakers, and is responsible for expressing coherence with preceding discourse. Topic therefore tends to appear at the beginning of a sentence. Comment, on the other hand, is usually new information, and tends to come near the end of the sentence.

The Topic-Comment Articulation Constraint

Werth expresses the requirement that Topic should come before Comment in terms of a global constraint on the arrangement of information, the Topic-Comment Articulation or TCA constraint. Rather than investigating the information structure of sentences directly to see if the constraint is observed, however, Werth postulates a correspondence between Topic and Comment and sentence accent. On Werth's account, Topic is often realised in speech as either unaccented (REDUCED or R in Werth's terminology) or with CONTRASTIVE (C) accent. The Comment, on the other hand, is ACCENTED (A) information. The constraint as a whole is phrased as follows [Werth 1984:220]:

The TCA Constraint: Semantic material is deployed in a discourse so as to respect the order: Anaphoric—Non-Anaphoric. This corresponds to the emphatic structure

R or C ... A

Corollary: to maintain this order in surface structure, syntactic elements may be moved, or a variant allowing the order may be preferred to one which loses it.

Werth's contention is that clefts of all kinds represent speakers' attempts to achieve an arrangement of information that satisfies the constraint.

In its strongest form, the TCA constraint makes the following predictions [Werth 1984:218ff]:

- R and C elements occur to the left
- A elements occur to the right

A sentence adhering to the TCA constraint (for example, one that can be analysed as following the schema R, C, A) has what Werth terms a STABLE structure; that is, a structure that is less likely to invite movement. If the pattern of the sentence is contrary to the TCA (e.g., if it follows the schema A, R, C), it is not stable and should encourage movement to a more stable arrangement. An intermediate pattern (which, presumably, may or may not be transformed) is one where the order of stresses are not actually

reversed (e.g. R, R, C). An example of each might be as follows: (2-39a) is a stable structure, (2-39b) is unstable, and (2-39c) is intermediate.

- (2-39) a What does John like?
Well, he likes BEANS.
R A
- b Who likes carrots?
JOHN likes BEANS.
A R C
- c And John likes beans?
No, John likes CARROTS.
R R C

The TCA Constraint and *It*-Clefts

The predictions the theory makes for clefts [Werth 1984:241] are the same as those for other sentences: we should expect to find an anaphoric leftward element (R or C) and an Accented rightward one. The constraint can therefore account for examples such as (2-40), which conforms more closely to the TCA constraint than the corresponding declarative in (2-41):

- (2-40) I know I should tell the vicar, but it's him that I'm aFRAID of.
R R A
- (2-41) I'm aFRAID of him.
R A R

Applying the Principle to Other Clefts

As Werth observes, the TCA principle does not account so well for clefts with other accentual structures. The cleft in (2-42), for example (Werth's 26a), has a Contrastive head and a Reduced complement, on Werth's analysis.

- (2-42) Though Johnson stepped up the war in Vietnam, it was KENNEDY who
C
first drew America into the conflict.
R

This sentence is not in direct conflict with the TCA constraint, but neither does it conform to it directly. Werth terms such sentences INTERMEDIATE structures.

This in itself is not a problem, but Werth notes that the declarative whose content would be equivalent to that of the cleft would itself have identical constituent and emphasis order: in effect, the movement has been from one intermediate structure to an identical one. In fact, then, clefts with this contrastive/reduced structure constitute

...*prima facie* evidence against the TCA constraint—evidently their function is different from the rearrangement of anaphoric material that TCA accounts for.

[Werth 1984:242]

Werth's constraint also runs into difficulties with *wh*-clefts. In a *wh*-cleft such as (2-5), where the *wh*-clause contains a contrastive element, both the *wh*-cleft and the corresponding declarative are intermediate structures. Example (2-43) (Werth's 34b) shows the TCA structure of the relevant portion of a declarative:

(2-43) He doesn't realise that ...
 R C R ...A

which, appearing as a *wh*-cleft, would receive the following analysis:

(2-44) What he doesn't realise is that ...
 R (R) R C

The comparison of the two examples shows a stable arrangement moving to another stable one. Although some of the examples Werth tests in each case move in a way predicted by the TCA constraint, many move between equivalent states. He notes that even this is often because the *wh*-element introducing the *wh*-cleft has been marked as anaphoric and therefore R; if this were not the case, the movement would appear to be from a more to a less stable arrangement—a situation as contradictory to the TCA as that of the *it*-cleft.

Werth therefore concludes:

These [examples] seem to suggest that, given an appropriate context, the TCA constraint cuts across whatever motivates the emphatic constructions, i.e. the two impulses are separate and perhaps independent.

[Werth 1984:242–243]

Werth goes on to argue that the emphatic constructions, clefts among them,

...are essentially RHETORICAL in nature, in that they are used to highlight semantic material, rather than simply redeploy it.

[Werth 1984:253]

Werth then attempts an alternative formulation to explain the behaviour of *it*-clefts, suggesting the function of the cleft to be that of demonstrating contrast on that head element. He notes, however, that in some cases there may be an extra contrastive emphasis in the complement. Werth interprets examples where two such points of contrast exist as demonstrating another motivation for clefting, akin to that suggested by Quirk *et al.*[1985], as we saw above: namely that of separating the two points of Contrast. The Contrastive element in the head 'often carries with it new Accented material', he claims, 'while the secondary contrast will usually be accompanied by repeated and Reduced material' [Werth 1984:254]. Note that this gives us an ordering exactly opposite to that predicted by the TCA; for example, (2-45) (Werth's example 31f [1984:245]) has the ordering A A C C R, rather than the predicted R C A:

(2-45) It's the *academic structure* of the UNIVERSITY that WE're concerned about.

A A C C R

Werth concludes that, for the *it*-cleft,

It is this reversal of the TCA order, we may assume, that gives the cleft its rhetorical effect.

[Werth 1984:254]

Discussion

Werth's attempt to salvage the TCA constraint by hypothesising that clefts serve to reverse it is not without its own problems. As we saw above, Werth was able to show that some clefts *did* conform to the TCA constraint, namely those which have anaphoric and/or unaccented head elements and new information in the complement. The reversal of TCA cannot explain these examples. *It*-clefts as a class, therefore, can neither be

described wholly in terms of an observation of TCA nor wholly in terms of a violation of it.

Werth's analysis, however, fares rather better for reverse *wh*-clefts. As we will see in chapter 5, the reverse *wh*-cleft head carries almost exclusively old information, and is either Reduced (i.e., unaccented) or Contrastive. New information, with corresponding accents, appears in the reverse *wh*-cleft complement. The common pattern for this kind of cleft under Werth's analysis would therefore be R (or C) followed by A, which is the pattern predicted by the TCA constraint.

However, because the declarative corresponding to any cleft may, as we have seen, be just as stable as the cleft, cleft constructions cannot be motivated solely on the basis of the requirement that Werth's TCA constraint—or any similar ordering principle—be satisfied. While we go on to show in chapter 5 that information-ordering principles are relevant in making the choice between clefts of various types, many other factors need to be considered (for example, presupposition, and the syntactic possibilities a sentence type offers). As it stands, Werth's account is unable to differentiate between clefts and any other construction that departs from canonical sentence order.

2.2.4 Dividing Given from New Information

In this section, we examine a claim that has a long history in the clefts literature, namely that cleft constructions serve the function of dividing given from new information, on some (explicitly or implicitly) Hallidayan notion of given and new. The earliest statement of this kind known to me is due to Prince:

although the *it*-cleft presents information (old vs. new) in aberrant order, it clearly marks which is which.

[Prince 1978:897]

A similar claim also appears in Geluykens [1984:68], Quirk *et al.* [1985:1387], and Cruttenden [1986:95]. Geluykens' version is as follows:

Clefted sentences are an ideal device to signal the Given-New distinction: they put the Given and New information in structurally very different parts

of the sentence.

The claim—or more commonly, the assumption—that given information appears exclusively in cleft complements, and new in cleft heads appears either implicitly or explicitly in Hornby [1972], Haviland and Clark [1974], Clark and Clark [1977:89], Glatt [1982:89], and Fletcher [1984]. However, it is contradicted by the data Prince [1978] and Geluykens [1984] present and by examples given already in this chapter, where new appears in complements and given in heads (for example, in Prince's *IP it*-cleft). As we have seen, Prince and Geluykens realise that given information can appear in cleft heads and new in cleft complements, but also that the complete opposite is possible. Their claim therefore relates to the *separation* of given and new information being the function of the cleft.

However, even this more general hypothesis is not supported by the evidence. First of all, it has been pointed out by Taglicht [1984] and others that the new information in the denotation of an accented constituent need not exhaust the information conveyed by that constituent. For example, in the *it*-cleft in (2-46), accent appears to indicate the new status of only part of the head constituent, namely *Bill*:

(2-46) It wasn't just John who was kicked out. It was John and BILL.

In cases such as this, the cleft head is composed of both new and given information. In fact, as Taglicht points out, it is possible to confine the scope of new information (described as 'focus' in Taglicht's terms) to a single semantic feature in the denotation of the constituent, resulting in a higher ratio of given to new information in the denotation of that constituent. In (2-47), new information is arguably confined to the feature [+FEMALE]:

(2-47) A: And then Billy Connolly came on?
B: No, next it was a comedienne that was on.

In (2-47), it is not at issue whether Billy Connolly is a comedian; that is, the contrast is not between Billy Connolly and a comedy act (this would require B to say *comedian* rather than *comedienne*). Instead, B makes a contrast between male and female comedians, the fact that both are comedy acts being given.



Not only is it possible to combine given and new information within the same syntactic part of the cleft, as (2-46) and (2-47) show, it is clearly the case that given or new can appear in both parts of the cleft at once. This appears to have been a conclusion that Quirk *et al.* [1985] in particular tried to avoid by their suggestion that, of the two nuclei they claim to be obligatory in cleft constructions, one—uncharacteristically for nuclei—had to mark given information. However, contrary to their claim, it is quite common for some new information to appear in both head and complement. *Wh*-clefts, for example, can have contrastive information (which falls under Halliday's definition of new) in the complement, and new information in the head, as in (2-48):

- (2-48) I'd rather they came later. What I DON'T want is for them to come while I'm at VINA's.

The problem of deciding on the scope of new information—that is, how much is new when the presence of new information is indicated by means of an accent—was first noted by Halliday [1967], and has appeared more recently in phonological approaches to focus (cf. Selkirk [1980], Gussenhoven [1983], Taglicht [1984]). It is generally held that the problem is not resolvable on an entirely syntactic basis, yet this is the force of what Prince, Geluykens, Quirk *et al.* and Cruttenden claim with their view of the cleft as providing a syntactic boundary between the two information types. While there may well be a correlation between syntactic structure and information structure, as we go on to examine in chapter five, the claim that a hard and fast syntactic division can be made between given and new information in clefts cannot be sustained.

2.2.5 Clefts and the Marking of Other Information Types: Conclusion

In this section, we have reviewed some claims regarding the kinds of information, other than focus, that clefts have been claimed to mark. We have also examined research that claims the function of clefts to be that of presenting information in particular ways, in terms of serial ordering and in terms of separating out different kinds of information.

Some problems that have occurred with the accounts discussed, in common with those occurring in relation to some of the focusing accounts discussed in the previous section, are based on too slender an acquaintance with the variety present in the data: this is true of Werth's suggestions regarding clefts as devices for presenting information according

to his definition of Topic-Comment ordering. In addition, our examination of accounts claiming that clefts divide new from given information revealed an underestimation of the varying scope of new information that is possible within constituents.

Prince's account, the only one that attempts to distinguish between *wh*-clefts and *it*-clefts, was found not to be correct as it stands, although her observations regarding the existence of different categories of *it*-clefts are important. In particular, the distinction she makes between GIVEN and KNOWN information was shown not to capture the fact that some *it*-clefts and *wh*-clefts carry the same kinds of information in their complements. In relation to this, we suggested that the label KNOWN represents a confusion of levels, being itself a side-effect arising in particular contexts, namely when new information is presented as a presupposition. We proposed as an alternative to Prince's distinction that *it*-clefts and *wh*-clefts should be seen as overlapping in the kinds of information they are able to convey as presupposition along some continuum of information types. This will be demonstrated and explained further in chapter 5.

2.3 An Overview of Previous Research: Summary

In spite of the variety of arguments put forward to explain the function of clefts and to characterise the kinds of information they appear to be able to convey, difficulties arising with the accounts suggested to date are often very similar.

The conclusions reached in this chapter can be summarised in the following five points:

1. On no definition of focus proposed so far in relation to clefts can the view of clefts as focusing constructions be substantiated.
2. Too little attention is paid to the variety of ACCENTUAL PATTERNS and INFORMATION STRUCTURES occurring in the data. The range available to clefts in both respects requires an extensive corpus-based study if general conclusions are to be drawn.
3. Possibly resulting from the above point, there is a tendency to incorrectly generalise about the correspondences that exist between accent and syntactic form, between information structure and syntactic form, and, where accent is taken into

account at all, between accent, information structure, and syntactic form.

4. In cases where an attempt is made to distinguish the different syntactic types of cleft, there is a tendency to assume a one-to-one mapping between the different syntactic structures and the kinds of information carried in each type.
5. Most strikingly, all the research examined attempts to explain the existence of cleft constructions in terms of a single level of description (information structure or focus, for example), neglecting important differences between clefts and other syntactic constructions that can be described at other levels. For example, almost all the accounts we have examined fail to mention the role of the cleft as a presuppositional construction, which differentiates it from other sentence types.

Because of this last point in particular, we advocate here a multi-level approach to the description of the function of clefts in discourse, where the distinctions between clefts and other syntactic constructions, and between the three types of cleft, can be explained in terms of their syntax, their semantics and their pragmatic function.

In the next chapter, we make the first step towards a multi-level model for the use of clefts in discourse: a study of the syntax of the three kinds of cleft construction.

Chapter 3

The Syntax of Cleft Constructions

While the nature of cleft constructions may seem intuitively clear, a full understanding of the range of data that fall within the definitions of the three cleft types, and the distinctions between them, cannot be reached without a detailed examination of the syntax of each of the structures involved. The particular concerns of this chapter are first of all to identify the structural distinctions between *it*-clefts, *wh*-clefts, and reverse *wh*-clefts that may be relevant considerations from the point of view of the speaker's choice of construction, and secondly to differentiate the clefts, or parts of them, from other kinds of construction that may be superficially similar.

Taking each cleft type in turn, we suggest an appropriate syntactic analysis, and look at how the analysis predicts the range of possible data. In the case of the *it*-cleft, the phrase-structure rules given by Gazdar *et al.* [1985] within the framework of Generalized Phrase Structure Grammar (GPSG) are adopted. This treatment is extended to *wh*-clefts and reverse *wh*-clefts, further GPSG rules being suggested to cope with each construction. For *wh*-clefts in particular, several competing analyses have been suggested: the distinctions between the analysis proposed here and these other analyses are pointed out and explained where appropriate. For all three cleft types, we are particularly interested in the kinds of constituent that can appear as heads, since these syntactic facts have effects with respect to the kinds of semantic object that cleft heads are able to specify. It seems that differing constraints operate on the instantiation of the head constituent in each type of cleft: we summarise the range of head constituent

types available to each cleft type at the end of the section in which that cleft type is discussed, and this aspect of the three cleft types is compared at the end of the chapter.

While the chapter concentrates on the formation of phrase-structure rules which, as is traditionally the aim of grammar, generate all and only the grammatical strings involved, it is also instructive to compare what is generated by the grammar with the frequency of occurrence of these strings in the corpora upon which much of this study is based. Some of the strings that the rules generate, and which are judged by informants to be grammatical, occur frequently in the data; others occur rarely or not at all. Some of the reasons for the scarcity of particular surface forms may be pragmatic, or connected with the content of what people decide to talk about. The approach we take to the problem is to construct phrase structure rules that generate the strings judged to be grammatical, but we also present the frequency data for comparison. In the suggestions for generating appropriate cleft constructions given in chapter 6, we propose that the tendencies towards particular surface strings can be captured at levels other than the syntactic level, either falling out of decisions made at the level of semantics or pragmatics, or adjusted for in terms of preference weightings.

The structure of the present chapter is as follows. In sections 3.1, 3.2, and 3.3, we discuss the syntax of the *it*-cleft, the *wh*-cleft, and the reverse *wh*-cleft respectively. Section 3.4 contains a discussion of the relevant features of the corpus data, while section 3.5 presents a summary of the findings of the chapter, paying particular attention to a comparison of the three types of cleft in terms of what syntactic constituents can appear as cleft head (defined below) in each case.

3.1 The *It*-Cleft

An example of the construction known as the IT-CLEFT appears in (3-1):

(3-1) It was William who was voted Greatest Cat.

As we noted in chapter 1, the practice throughout this thesis will be to refer to the immediate complement of the copula in each kind of cleft—in this case the NP *William*—as the HEAD of the cleft. The clausal constituent that comprises the remainder of the

cleft, in this case *who was voted Greatest Cat*, will be termed the cleft COMPLEMENT (or, as we observed for *wh*-clefts, the WH-CLAUSE).

3.1.1 Phrase-Structure Rules for *it*-clefts

In our description of *it*-clefts, we will be adopting the phrase-structure rules developed in Generalized Phrase Structure Grammar (henceforward GPSG; cf. Gazdar *et al.* [1985]). Gazdar *et al.* [1985:159ff] give two rules which in combination are intended to cover the range of the clefts data. The rules are as follows:

$$(3-2) \quad VP[+ it] \longrightarrow H[44], NP, S [+ R]$$

$$(3-3) \quad VP[+ it] \longrightarrow H[44], x^2, S [FIN]/x^2$$

In each rule, the category $H[44]$ stands for the main verb *be*; the category x^2 in rule (3-3) can be characterised as representing any major phrasal constituent (such as NP, AP, PP), while the constituent $S[FIN]/x^2$ stands for a finite sentential constituent containing a gap of category x^2 . That is, it corresponds to an *s* which, in transformational terms, has had a constituent of category x^2 extracted from it. However, in the framework of GPSG, there are no 'extractions' or 'movements' as such. Rather, there is a systematic use of features for encoding UNBOUNDED DEPENDENCIES between fillers and gaps. For the sake of convenience, the term 'extraction' will be used from time to time during the discussion to follow. This is not, however, intended to imply a transformational analysis.

We will discuss the range of data covered by each of the two rules in turn. An example of the tree structure associated with rule (3-2) appears in figure 3.1.

The Relative-Clause Rule

The first point to note regarding rule (3-2) is that it generates only NP-headed *it*-clefts, rather than the full range of *it*-cleft head types. Secondly, the rule generates a normal restrictive relative clause as the complement of the cleft. On this analysis, the relativiser involved plays the role of a subject or object NP argument in the embedded sentential

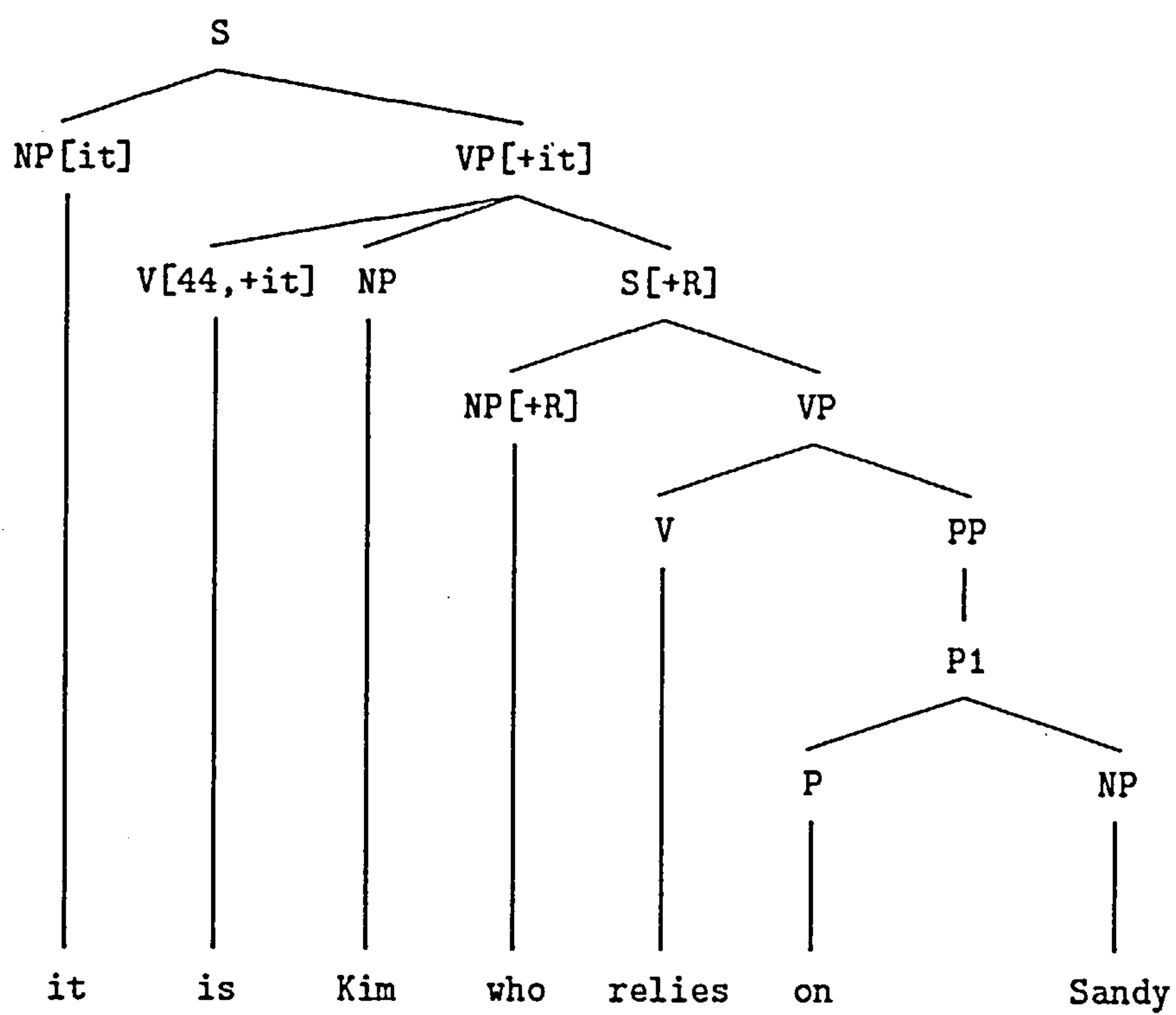


Figure 3.1: Phrase-marker for an *it*-cleft generated by rule (3-2)

constituent. Two important features fall out of this view of the role of the relativiser: one is a correct account of person agreement, and we discuss this later in this section. The second is that, as is common to the analysis of both *it*-cleft complements and relative clauses in GPSG, the relativiser cannot be omitted when it plays the role of subject in the embedded clause. The fact that this feature appears in both types of construction is also noted by Sornicola [1988] and Quirk *et al.* [1985]. In basing the (3-2) analysis of *it*-clefts on that of restrictive relative clauses, Gazdar *et al.*'s rules correctly predict that *it*-cleft complement relativisers that are covered by this rule can only be omitted when they act as object in the embedded clause. Relative clause examples are given in (3-4); *it*-cleft examples in (3-5). In the (a) example in each case, the relativiser is object; in the (b) cases, it is subject:

- (3-4) a The man I saw has come back.
 b *The man saw me has come back.

- (3-5) a It was the man I saw first.
 b *It was the man saw me first.

The rule also predicts that *it*-clefts with NP heads can be generated with the full range of appropriate relative-clause relativisers, such as *that*, *which*, *who*, *whom* and compounds of these items with prepositions such as *on whom* and *with which*.

Note that this prediction is incompatible with a conclusion of Quirk *et al.* [1985], who argue that '*whom* and *which* are only marginally possible [in *it*-clefts], and it is virtually impossible to use *whom* or *which* preceded by a preposition' [Quirk *et al.* 1985:1397]. This position seems to be difficult to substantiate, since their example, given here in (3-6), which they claim only admits a non-cleft reading, seems to be equally appropriate in the context given in (3-7), in which a cleft reading is induced:

- (3-6) It was the dog to which I gave the water.

- (3-7) It was the dog to which I gave the water, not the cat.

Examples generated by rule (3-2) include the following (cf. Gazdar *et al.* [1985:158]):

- (3-8) a It is Kim on whom Sandy relies.
 b It is Kim who Sandy relies on.
 c It is Kim that relies on Sandy.

The analogy between the complements of NP-headed *it*-clefts and the relative clause is further supported by the fact that PIED-PIPING (cf. Ross [1967]) occurs in these *it*-clefts just as it does in relative clauses. The term developed from the view that relativisation, seen in some abstract sense, involves the FRONTING of a *wh*-constituent, a progression that is represented schematically in (3-9). The argument runs that when such fronting has taken place, there is a choice between leaving associated prepositions STRANDED at the end of the clause, as in (3-10), or pied-piping them, like the piper of Hamelin, along to a position accompanying the relativiser, as in (3-11):

- (3-9) a I spoke to the man.
b The man I spoke to who.
c The man who I spoke to.

(3-10) The man who I spoke to.

(3-11) The man to whom I spoke.

The acceptability of both stranded and pied-piped prepositions in relative clauses is identical in the case of *it*-clefts, as follows:

(3-12) It was the man who I spoke to.

(3-13) It was the man to whom I spoke.

It will be noted that the phrase structure given by rule (3-2) does not presuppose that the head and relative clause form a constituent—such as *Kim who relies on Sandy*—that can be analysed as NP. Although the NP analysis is a common one for relative clauses and their heads, it is clear that the relative clause of the *it*-cleft does not form a constituent with the cleft head. This is demonstrated by examples such as (3-14), where an attempt is made to extract the resulting ‘constituent’:

(3-14) *He said it was John that’s an interesting guy, and John that’s an interesting guy it is.

The structure given by rule (3-2) predicts that the extraction will be unacceptable.

Rule (3-2), although capturing a range of *it*-cleft behaviour, does not cater for all the data, however. For example, Sornicola [1988] notes that restrictive relative clauses

cannot normally modify proper nouns:

- (3-15) a Which company are you going to ring?
b *I'll ring Momma's who make the best pizzas.

In addition, rule (3-2) does not cover the many cases where *it*-clefts have heads that are not NPs. In these cases, as Quirk *et al.* [1985:1387] have observed, the relativiser *that* does not function strictly as a pronoun, as rule (3-2) requires, since it can also occur in sentences where it serves as a pro-form for the non-nominal constituent. For example, in (3-16), *that* substitutes for the reason adverbial *because he was ill*:

- (3-16) It was because he was ill that we decided to return.

Quirk *et al.* note that while *that* is adaptable to such non-pronominal use, *which* is not so adaptable, since it cannot be used in contexts in which the non-pronominal function is required:

- (3-17) *It was because he was ill which we decided to return.

We would therefore expect *which* to appear only with NP-headed clefts, and this is borne out by the data. Gazdar *et al.*'s rule (3-3) deals with these non-pronominal cases.

The Complementised Sentence Rule

Rule (3-3) allows any constituent that can be extracted from *s* to be cleft head, leaving the cleft complement as a sentential constituent containing a dependent gap, rather than a relative clause. The phrase-marker given by Gazdar *et al.* [1985:159] for an example that the rule would generate appears in figure 3.2.

First of all, as we noted above, only the sentential complementiser *that* appears in examples covered by rule (3-3). The rule therefore excludes examples such as the following:

- (3-18) *It was on Kim on whom Sandy relied.

- (3-19) *It was to Kim who John gave a book.

The range of constituents that can act as XP, i.e. that can be extracted from *s* are,

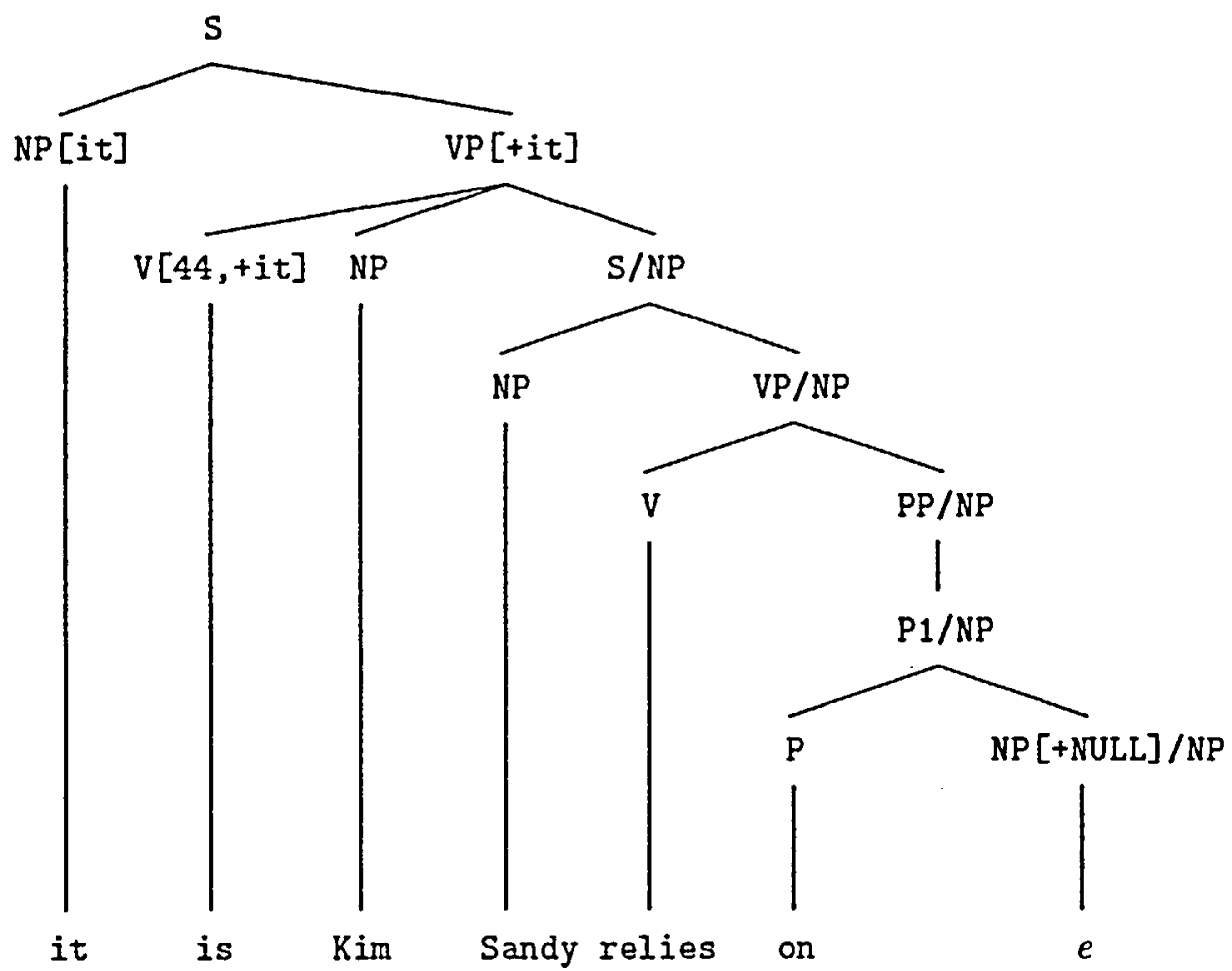


Figure 3.2: Phrase-marker for *it*-cleft generated by rule (3-3)

in GPSG as elsewhere, taken to be any major phrasal constituent, such as PP, AP, VP or NP. XP cannot, however, be instantiated by a particle, NP with possessive marking, non-finite VP, V, or A. In fact, the constituents predicted to be acceptable as XP are the same as those that, on similar rules, can be extracted in sentential constructions other than *it*-clefts, such as topicalisation and *wh*-fronting. We shall see below, however, that the possible instantiations of XP in *it*-clefts are not identical either to the constituents possible in topicalisation and *wh*-fronting, or to those possible for other types of cleft. It is not to the purpose here to attempt to adduce rules for the instantiation of XP in each case, and so in what follows, we will simply draw attention to the unacceptable cases, leaving the particular constraints that operate in each case as a topic for further research.

Rule (3-3) predicts that clefts with heads that are not major phrasal constituents will not be acceptable. It therefore makes the correct predictions for examples such as the following. In (3-20a), a possessive NP has been extracted; in (3-20b), a particle, and in (3-20c), a non-finite VP (in this particular case, a bare infinitive):

- (3-20) a *It was Kim's I read a book of.
 b *It was up he called John.
 c *It is be beaten by a six-year-old that I couldn't bear.

The rules also predict that cases in which the head is a non-extracted constituent will not be acceptable. This applies to the case of sentential adverbs, which, although they can appear at the in various positions throughout the sentence, do not create dependencies in the normal sense since it is never clear whether extraction has taken place at all. The explanation of such constituents as non-extracted would account for the unacceptability of (3-21):

- (3-21) *It is frankly that I think he's a nuisance.

While VP is predicted to be one of the major constituents that can be extracted from S, not all kinds of VP can be cleft head. First of all, in all three kinds of cleft, finite VPs cannot be cleft head:

- (3-22) a John bit me.
 b *It was bit me that John.

Extraction of a range of non-finite VPs is possible, however. Purposive *to*-infinitive clauses, such as that forming cleft head in (3-23a), non-finite VPs in *-ing* acting as subject, as in (3-23b), and passive participle clauses, such as (3-23c), are all possible:

- (3-23) a It was to get some shaving cream that Bill set out for the store.
b It was making all those beds that was the worst thing.
c It was jammed under the sofa that I found your wallet.

Non-purposive *to*-infinitive clauses such as that appearing as head of (3-24a), *-ing* clauses acting as complement of an embedded VP such as that in (3-24b), and bare infinitives such as that in (3-24c), are all ungrammatical, however:

- (3-24) a *It's to go out that I'd prefer.
b *It's making the beds that John is [doing].
c *It's make the beds that I'll have to.

Note that in (3-24b), *do*-insertion appears unable to support the VP in the cleft head. While *do*-insertion in *wh*-clefts allows a range of non-finite VPs in head position as we shall see in section 3.2, neither progressive non-finite VPs, such as that in (3-24b), nor non-progressive non-finite VPs such as that in (3-25), appear to be supported by *do* in the case of the *it*-cleft:

- (3-25) *It is make the beds that John does.

Unlike reverse *wh*-clefts, *it*-clefts are able to take phrases with *not only* and *not just* as head. In addition, uniquely among the cleft constructions, they are able to take phrases with *not until*.

- (3-26) a It is not only/not just the wind that makes the noise.
b What makes the noise is not only/not just the wind.
c *Not only/not just the wind is what makes the noise.

- (3-27) a It was not until three that she left.
b *When she left was not until three.
c *Not until three was when she left.

Two final constraints on the instantiation of XP are as follows. First, it appears that XP for *it*-clefts cannot be either a universally quantified phrase such as (3-28):

- (3-28) *It is every boy that Mary likes.

Nor, as Halvorsen [1978] points out, can it be an NP with a negative polarity item, such as (3-29):

(3-29) *It is any marbles that John doesn't want.

Second, although this is strictly a pragmatic issue, it is worth noting here that certain anaphoric NPs, namely *that* and *it*, are less than acceptable as head of an *it*-cleft. *It* is completely unacceptable, while *that* is merely marginal:

(3-30) *It was it that I was talking about.

(3-31) ?It was that that I was talking about.

Interestingly, *it* cannot appear as head of a topicalised clause either, while *that* can:

(3-32) *I looked for the spanner. It I could carry.

(3-33) I looked for the spanner. That I could carry.

This pragmatic constraint on the instantiation of XP is discussed further in chapter 5.

3.1.2 *It*-clefts and Topicalisation

The most interesting cases are those where the behaviour of the *it*-cleft appears to diverge from that of the putatively analogous topicalised and *wh*-fronted constructions. In some cases, pragmatic principles can be brought to bear to explain the distinctions involved; in others, the reasons for the divergence are not so clear. We will note three particular cases where *it*-cleft behaviour does differ, only the third of which appears to have a pragmatic explanation. The cases are as follows:

- AP cannot always act as head of an *it*-cleft, although it can in topicalised clauses.
- As the rules for *it*-clefts predict, ADV cannot appear as a cleft head; it can, however, appear as the fronted constituent of a topicalised clause.
- Sentential complements of non-factive verbs cannot be *it*-cleft heads, although this does not appear to be a restriction on topicalised clauses

Taking the AP issue first, Gazdar *et al.* [1985:159] note that only in some dialects can AP be extracted, their example being as follows:

(3-34) % It is very enthusiastic that Sean is.

In *it*-clefts, AP appears in fact to be acceptable when it appears in an adjunctive role modifying the main verb, as in (3-35). In this example, the action of the main verb could be described as *painting green*:

(3-35) It was lurid green that he painted his boat

However, AP is either marginal or unacceptable when it appears as complement of an intensive verb:

- (3-36) a ?It was absolutely delicious that your pie was.
b ?It was morose that he became.
c *It was furry that it felt.
d *It is cacophonous that this sounds.

Interestingly, extraction of the same kind as that shown in (3-36) is acceptable in topicalised clauses:

- (3-37) a He said your pie would be absolutely delicious, and absolutely delicious it was.
b They said John was morose, and morose he is.

A related construction is *though* fronting, which permits AP in cases where *it*-clefts do not:

- (3-38) a Short though he is, he made it into the basketball team.
b *It is short that he is.

It has been suggested that only *nounlike* APs can be clefted, which might explain the acceptability of (3-34). However, why this constraint should operate on *it*-clefts and not on the corresponding topicalised clauses is not immediately clear.

A second divergence from the case of topicalisation occurs in the case of ADV as cleft head. *It*-clefts appear to be unable to take *-ly* adverbs in particular as cleft head, as is demonstrated in (3-39). If the adverb is analysed as simply A, rather than as a

realisation of ADVP, this unacceptability is predicted by rule (3-3):

- (3-39) a *It was tenderly he sang.
b Tenderly he sang.

ADVP, on the other hand, is acceptable, again as the rules predict:

- (3-40) It was with great tenderness that he sang.

The third and final case in which topicalisation and *it*-clefts differ is in the acceptability of sentential complements of factive and non-factive verbs. As Delahunty [1984:82] notes, the complements of non-factives such as *think*, *write* and *say* cannot appear as the heads of *it*-clefts, while those of factives such as *forget*, *remember* and *regret* can be. The following example demonstrates:

- (3-41) a *It was that he was coming home that he wrote.
b It was that he was coming home that he regretted.

Sentential complements of both non-factive and factive verbs can be topicalised, however:

- (3-42) a That he was coming home he regretted, but his decision was final.
b That he was coming home he wrote a week ago; why he has not appeared is a mystery.

The constraint on the sentential complements of non-factives appearing as the heads of *it*-clefts is not necessarily exclusive to clefts, since a plausible and more general explanation is available on the basis of pragmatic principles concerning the use of such complements. Factive verbs are so called because their complements have the status of fact: the complement is presupposed. This makes them particularly well-suited to appearing as cleft heads, since, as we show in chapter four of this thesis, cleft heads act in a manner very similar to that of definite referring expressions in discourse. Their similarity lies in the fact that an antecedent seems to be required for them both in the discourse, in each case because of an existential presupposition that arises from syntactic structure. In the case of definite referring expressions the presupposition arises from the determiner *the*; in the case of clefts, it arises from the presence of an existentially-quantified variable that appears in the presupposition conveyed by the cleft complement. This is discussed fully in chapter four: for the moment, it is sufficient to note that the

complements of factive verbs have the required quasi-definite status in the discourse to appear appropriately as cleft heads, while those of non-factives do not.

While the reasons behind the three points of divergence between *it*-clefts and other constructions where similar extraction takes place are not always clear-cut, the above evidence demonstrates that slightly different constraints operate on what constituent type XP can be for the purposes of cleft rule (3-3) than might operate on a similar topicalisation or *wh*-fronting rule of the same type. We will summarise what the constraints on XP seem to be at the end of the section.

3.1.3 Agreement in *It*-clefts

Before we go on to look at the constraints operating on the extraction of constituents from *s* in *wh*-clefts and reverse *wh*-clefts, we should look at the behaviour of *it*-clefts with respect to agreement, in particular, agreement between the main verb embedded in the complement and the cleft copula. First of all, as Halvorsen [1978] and Sornicola [1988] point out (following Jespersen [1949]), the *it*-cleft copula no longer agrees with the cleft head, as it did in Middle English. An oft-quoted example from Chaucer demonstrates this agreement:

(3-43) It am I that loveth so hote Emilye the brighte.

Copular agreement of this nature is generally judged to be ungrammatical in Modern English:

(3-44) *It am I that is going to the pictures tonight.

(3-45) *It were you lot that had the Radio Times last.

The agreement patterns between the verb embedded in the cleft complement and the copula are rather less simple. When the head of the cleft is object of the embedded verb, agreement is straightforwardly between the embedded verb and grammatical subject of the embedded sentence, as might be expected:

(3-46) It is me that Mary talks to most.

Agreement of the verb in the complement clause is not so straightforward when the subject of that verb appears as cleft head. Halvorsen [1978:51] notes that, in examples such as (3-47), the verb in the *it*-cleft complement agrees with the subject cleft head in number, but not in person:

(3-47) It's me who is/*am responsible.

Akmajian [1970:153] and Huddleston [1984:462], however, give examples where both person and number agreement seem to be taking place between the verb and the *it*-cleft head. Akmajian [1970] points out that, in dialects where such examples are acceptable, when the head of the cleft is a nominative pronoun, person and number agreement both occur:¹

- (3-48) a It is I who am/*is responsible.
b It is the boys who are responsible.

Sornicola [1988] proposes that, while number agreement is always controlled by the cleft head, person agreement only appears if the head takes the nominative case. The contrast between (3-48a), where the cleft head is nominative, and (3-46), where it is not, shows this to be the case. Note first of all that the examples we are discussing are those that would be generated by rule (3-2) given above, in which the cleft complement is analysed as a relative clause. Rule (3-2), in the cases we are concerned with², casts the relative pronoun as grammatical subject of the complement. On the GPSG analysis, we could postulate that the verb agrees by default with the relative pronoun. We can therefore describe this default agreement simply as agreement with the subject. Because the element agreed with—the subject relative pronoun—exists only in third person form, default agreement is third person. In cases where the cleft head is deliberately marked as subject by the use of the nominative case, however, as occurs in (3-48a), we can say that this element is being marked as the dominant subject element. Agreement in these cases defaults to that element, and not to the relative pronoun³.

¹In Italian, agreement with the head NP is unvarying; in French, the cleft shows variations in the extent of this agreement, particularly in older rural varieties. For a full discussion, see Sornicola [1988].

²Nominative marking cannot appear in the cleft head when the relative pronoun is not acting as subject: consider, for example, **It is I that Mark hates*.

³Note, however, that in the Middle English example in (3-42) agreement does not default to this element even though the cleft head is marked as nominative. Agreement instead appears to be between the embedded verb (e.g. *loveth*) and the third-person relative pronoun in all cases.

3.1.4 Summary

In this section, two GPSG rules for generating *it*-clefts were adopted. Rule (3-2) describes NP-headed *it*-clefts with relative clause complements; rule (3-3) caters for *it*-clefts with heads of constituent type XP, whose complement is a sentential constituent containing a gap of category XP. The constraints on what constituent types could instantiate XP were examined, and it was demonstrated that all major phrasal constituents could appear except for finite VP, non-purposive *to*-infinitive, non-finite *-ing* VP as complement of embedded VP, universally quantified or negative polarity NP, some APs, *-ly* adverbs, sentential complements of non-factive verbs, sentential adverbs, and bare infinitive clauses. It was also shown that *it*-clefts cannot have *it* as head, while *that* is marginal, but possible.

As far as agreement of the main verb of the complement clause is concerned, we concluded that number agreement is always with the cleft head, while person agreement seems to default to the relative pronoun unless the cleft head is marked as nominative. In these cases, person agreement is with the head element.

3.2 The *Wh*-Cleft

The copular construction known as the *wh*-cleft (sometimes referred to as the PSEUDO-CLEFT) construction is exemplified in (3-49):

(3-49) What Andrew wants most is to find a nice cover for his book.

As is the practice throughout this thesis, in the discussion to follow, the relative-like clause that introduces the *wh*-cleft will be termed the WH-CLAUSE (in (3-49), *what Andrew wants most*) and the immediate complement of the copular main verb will be termed the HEAD of the cleft (in (3-49), *to find a nice cover for his book*).

3.2.1 Phrase-Structure Rules for *Wh*-clefts

It is easy to see that the *wh*-clause of the *wh*-cleft contains a gap, and that this may be embedded arbitrarily deeply. For example:

- (3-50) What Andrew seems to have persuaded Sue to want to get
Alex to do ...

What is less clear is what constitutes the filler corresponding to the gap. One hypothesis is that the initial *wh*-element, such as *what* in the examples in (3-51), and *where* in the examples in (3-52), plays this role:

- (3-51) a What_{NP} Andrew wants most *e*_{NP}.
b *What Andrew wants to put the car *e*.
(3-52) a Where_{PP} Andrew wants to put the car *e*_{PP}.
b *Where Andrew wants most *e*.

In (3-52a), the *where* is presumably a *wh*-filler of category PP which enters into a dependency with the PP gap after *the car*. In cases where category of the *wh*-filler and that of the gap do not match, such as in (3-50b) and (3-52b), the sentence is ungrammatical.

Although there is clearly a dependency of some sort between the *wh*-filler of the *wh*-clause and the head of the whole *wh*-cleft, it is difficult to subsume it under the general mechanisms of filler-gap dependencies using GPSG-style slash features, unlike the dependency within the *wh*-clause just discussed. If such a gap-filler dependency did exist, the analysis of the *wh*-cleft might be represented by a rule such as (3-53b):

- (3-53) a What John wants is beans.
b $s \longrightarrow s/XP \vee XP$

However, on this analysis, we would expect it to be possible to topicalise the filler constituent XP, as follows:

- (3-54) a *Beans, what John wants is
b $s \longrightarrow XP[TOP] s/XP \vee$

The lack of acceptability of this topicalisation suggests that the relationship between the cleft head constituent and the *wh*-clause is something other than a straightforward

gap-filler dependency. One alternative possibility is that the dependency is purely semantic. A second, and perhaps more appealing, suggestion is that we could capture it syntactically by postulating a category-valued foot feature, say CL, which is analogous to SLASH, but appears only in *wh*-clefts.⁴ That is, consider the following rule: Jacobson discontinuous 1987 .]

$$(3-55) \quad s \longrightarrow s[CL \text{ } XP] \text{ } v[106] \text{ } XP$$

We discuss later in this section the reasons behind this flat structure; for the moment note that the specification [CL XP] is intended to encode the dependency between the XP category of the head and the XP category of the *wh*-element in the *wh*-clause.

In order to get this to work, it seems that a second rule of the following kind might be necessary:

$$(3-56) \quad s[CL \text{ } XP] \longrightarrow XP[WH \text{ } XP] \text{ } s/XP$$

However, let us suppose that lexical entries of the following sort are available:

- (3-57) a *what* NP[CL NP]
 b *where* PP[CL PP]
 c *who* NP[CL NP]

We could now replace (3-56) by (3-58), which draws upon lexical entries such as those in (3-57):

$$(3-58) \quad s[CL \text{ } XP] \longrightarrow XP[CL \text{ } XP] \text{ } s/XP$$

The conventions which govern foot features in GPSG can be used to ensure that the [CL XP] on the mother of (3-58) will also occur on one of the daughters. If we assume, in addition, that [CL XP] is prevented from co-occurring with [SLASH XP] by a Feature Co-occurrence Restriction (FCR), then (3-58) does not need to be separately written in the grammar, but will follow from the Foot Feature Principle (FFP; cf. Gazdar *et al.* [1985]) and appropriate FCRs.

⁴Similar proposals for a slash-type feature which is not unbounded have been made in the GPSG literature, though for somewhat different purposes: see, for example, Jacobson [1987].

It is not our intention to look for detailed arguments in favour of one approach or another, since we are not primarily concerned with this aspect of the syntax-semantics interaction in this thesis. For convenience, we shall just adopt the [CL XP] analysis in the following discussion. Later in this section, we use this analysis as the basis of a comparison between the *wh*-cleft and other kinds of *wh*-construction.

As in the discussion of *it*-clefts, we are interested in examining what constituent types can instantiate XP, as well as in demonstrating that the analysis given in the rule is appropriate for the construction in question. We will make some remarks concerning the comparison of *wh*-clefts with *it*-clefts and reverse *wh*-clefts during the course of the following section, but will provide a thorough comparison of the three construction types in section 3.5.

3.2.2 The Instantiation of XP

In the case of the *wh*-cleft, a broader range of constituents can act as XP—i.e. as the head of the cleft—than can appear as head of *it*-clefts. However, the constraint that XP has to be a major phrasal constituent still holds, predicting that the examples below will not be grammatical *wh*-clefts. In (3-59), the cleft head is a preposition extracted from a phrasal VP; in (3-60) the head is a possessive noun extracted from a possessive NP, and in (3-61), the head is formed from a concatenation of two unrelated constituents:

(3-59) *What/how he called John was up.

(3-60) *Who he read a book of was Kim's.

(3-61) *Who/What they are making is Mike captain.

Several constraints on XP do arise, and these are as follows. First of all, we saw in the case of *it*-clefts that AP heads were only marginally acceptable. Difficulties also arise with this category of head in *wh*-clefts, but the problems are somewhat different in nature. Often, when an AP is placed as head of a *wh*-cleft, instead of the result being ungrammatical, a non-cleft reading arises. This reading results from a different semantic and syntactic analysis of the same string. In the non-cleft case, the *wh*-clause is interpreted as a referential free relative clause, while the complement of the copula is

interpreted as predicating some property of the object referred to by the free relative clause. Following Higgins [1979], we term this a PREDICATIONAL reading. For example, in (3-62), the *wh*-clause *what Lisa is chasing* refers to an entity, and the predicative complement *covered in scales* predicates a property of that entity:

(3-62) What Lisa is chasing is covered in scales.

Although ungrammaticality is also an issue in certain cases, the main issue to take note of in the discussion of AP heads—as in predicative heads in general—is to make sure that non-cleft, predicative readings do not occur. Broadly speaking, AP heads do not result in a predicative reading in *wh*-clefts as long as the main verb of the *wh*-clause is INTENSIVE: that is, if it is a copula or a member of the class of verbs (such as *look* and *appear*) in which the complement predicates a property of the entity referred to by the subject phrase. A cleft reading can therefore be preserved in examples such as (3-63), where the main verb of the *wh*-clause is a copula:

(3-63) What John is is interesting.

Where the *wh*-clause main verb is non-intensive, an unambiguously predication reading occurs:

(3-64) What John does is interesting.

Examples of other intensive verbs which allow the cleft reading to be preserved in the presence of an AP or A head are *become*, *feel*, and *sound*, exemplified as follows. As we noted in the previous section, the corresponding *it*-cleft is not acceptable in any of these cases:

(3-65) a What he became was morose.
b *It was morose that he became.

(3-66) a How it felt was furry.
b *It was furry that it felt.

(3-67) a How this sounds is cacophonous.
b *It is cacophonous that this sounds.

There are intensive verbs, however, which do not permit either a cleft or a predication reading of the sentence:

(3-68) ?What John seems is unhappy.

(3-69) ?What he remained was morose.

Turning to the acceptability of non-finite VP as cleft head, we saw in the case of the *it*-cleft that three types of phrase were possible: purposive *to*-infinitives, *-ing* participle phrases when used as subject, and passive participle phrases. All three are also possible for *wh*-clefts, examples of each type of head appearing as follows:

(3-70) What Bill set out for the store for was to get shaving cream.

(3-71) What was the worst was making all those beds.

(3-72) Where I found your wallet was jammed under the settee.

However, in addition to these types, a further range is also possible for *wh*-clefts. Non-purposive *to*-infinitives are possible:

(3-73) a What I'd prefer is to get a more appley green.

b *It's to get a more appley green that I'd prefer.

More possibilities are opened up by the ability of the *wh*-cleft to allow *do*-insertion in the *wh*-clause. This allows bare infinitives to appear as cleft head:

(3-74) a What she would like to do is leave straight away.

b *It's leave straight away that she would like to do.

We saw in the previous section that participle phrases with *-ing*, while acceptable as subject, are not acceptable as predicative complements of *be* in the *it*-cleft. They are, however, acceptable as *wh*-cleft heads, again because *do* can appear in the *wh*-clause. Examples of both *it*-cleft and *wh*-cleft appear in (3-75):

(3-75) a *It is making the beds that John is [doing].

b What John is doing is making the beds.

Do does not, however, allow non-finite VPs of all kinds to appear as *wh*-cleft head. In particular, it cannot allow stative, non-volitional verbs such as *suffer*, *exist* and non-

volitional meanings of *see* to appear as cleft head, because binding with *do* requires volition on the part of the agent of the action. We can see this through the contrast of volitional and non-volitional uses of *see*, in the following example: *see a psychiatrist* is volitional, while *see a stone* is not:

- (3-76) a What Mary does is see a psychiatrist.
b *What Mary does is see a stone.

As we noted earlier, support of non-finite VPs through the use of *do* is not possible for *it*-clefts.

A final point to note on non-finite VPs as heads is that non-finite *-ing* clauses may be extracted out of PP to form the head of the cleft, but only if the PP has an adjectival function, as in (3-77), rather than an adverbial one, as in (3-78):

- (3-77) What he was hopeless at was writing letters.

- (3-78) *What Mary disposed of the car by was driving it over a cliff.

We saw in the case of *it*-clefts that there are restrictions on the appearance of ADVP as head, in particular that *-ly* adverbs were only marginally acceptable. Restrictions on the appearance of various kinds of ADVP in head position of a *wh*-cleft or reverse *wh*-cleft are of rather a different sort. The constraints can be couched in terms of a distinction among different kinds of PROCESS adverbials made by Quirk *et al.* [1985]. The three subtypes of these adverbials that interest us are what Quirk *et al.* term MANNER adverbials, such as *casually*, *with deference* and *carefully*, MEANS adverbials, such as *by bus* and *through insight*, and INSTRUMENT adverbials, such as *with a crowbar*. The key to the constraint on the appearance of these adverbials as *wh*-cleft head is that, of the three categories, only means and instrument adverbials can be bound by the *wh*-word *how*, while the appearance of *how* with manner adverbials is only marginally acceptable:

- (3-79) a ?How he did it was with deference.
b How he got there was by bus.
c How he broke it open was with a crowbar.

The reason behind this unacceptability seems to be that *how* can bind only adverbials that can be characterised as VOLITIONAL. The less volitional the adverb, the less ac-

ceptable the *wh*-cleft with *how*. Volitional adverbials such as the means and instrument examples above are perfectly acceptable; manner adverbials such as the head of (3-79a) could be described as semi-volitional, resulting in a marginally acceptable reading, while the definitively non-volitional manner adverbials, many of them (but not all) *-ly* adverbs, are definitely unacceptable. Examples of this latter kind appear in (3-80):

- (3-80) a *How he did it was luckily.
 b *How he won was by chance.
 c *How he smiled was happily.

A second restriction on the appearance of certain types of ADVP also seems to be based on the lack of appropriate *wh*-words to bind them in a *wh*-cleft construction. As in the case of non-volitional manner adverbials, indexical adverbials expressing time (such as *immediately*, *first*, etc.) appear not to be correctly bound by any of the available *wh*-words. Examples are as follows:

- (3-81) *How/*When she came in was immediately.

- (3-82) *How/*When she came in was first.

We noted in the case of *it*-clefts with sentential complements as heads that only the complements of factive verbs such as *regret* could appear as cleft head. For *wh*-clefts, the complements of both factive and non-factive verbs can appear, as follows:

- (3-83) a What he regretted was that he was coming home.
 b It was that he was coming home that he regretted.

- (3-84) a What he wrote was that he was coming home.
 b *It was that he was coming home that he wrote.

We saw in the case of the *it*-cleft that quantifier phrases are unacceptable as head, as was NP with a negative polarity item. This is the case for the reverse *wh*-cleft also, as we will see in the next section. *Wh*-clefts, however, are the exception: as Halvorsen [1978] points out, *wh*-clefts are able to take both kinds of expression as head. (3-85a) demonstrates a universally quantified phrase, (3-85b) an NP with negative polarity determiner:

- (3-85) a What John has seen is every film of Coppola's
 b What John doesn't want are any marbles.

A final point to note regarding the heads of *wh*-clefts is that the head constituent cannot be anaphoric⁵:

(3-87) *What I want is it.

(3-88) *What I want is that.

(3-89) *What I want is this.

(3-90) *Where I want to go is there.

Since this constraint is related to pragmatic considerations, it will be discussed further in chapter 5.

3.2.3 The Analysis of *Wh*-clefts

The key points of the analysis of *wh*-clefts argued for here are as follows:

- the *wh*-clause is not analysed as NP, but as S; and
- the copula and its complement—the cleft head—are not analysed as VP, but as two constituents, V and XP.

We wish also to argue for a syntactic distinction between the *wh*-cleft and the *wh*-relative construction, in contrast to the conclusions of Halvorsen [1978], Huddleston [1984], and Quirk *et al.* [1985], who analyse the *wh*-cleft as containing a free relative clause. In addition, we wish to point out distinctions between both clause types and *wh*-interrogatives, since it has been claimed (cf. Faraci [1971] and Nakada [1973]) that the *wh*-cleft *wh*-clause is an interrogative. In the discussion to follow, we will refer to *wh*-cleft *wh*-clauses as WHC, *wh*-relative clauses as WHREL, and *wh*-interrogative clauses as WHI.

⁵It has been suggested to me that the unacceptability of these examples is not due to the presence of an anaphoric item but to lack of 'heaviness' on the head constituent. This seems unlikely, since the available deictic readings of the same lexical items are acceptable. For example:

(3-86) Headache? What you need is THIS! (*holds up packet of aspirin*)

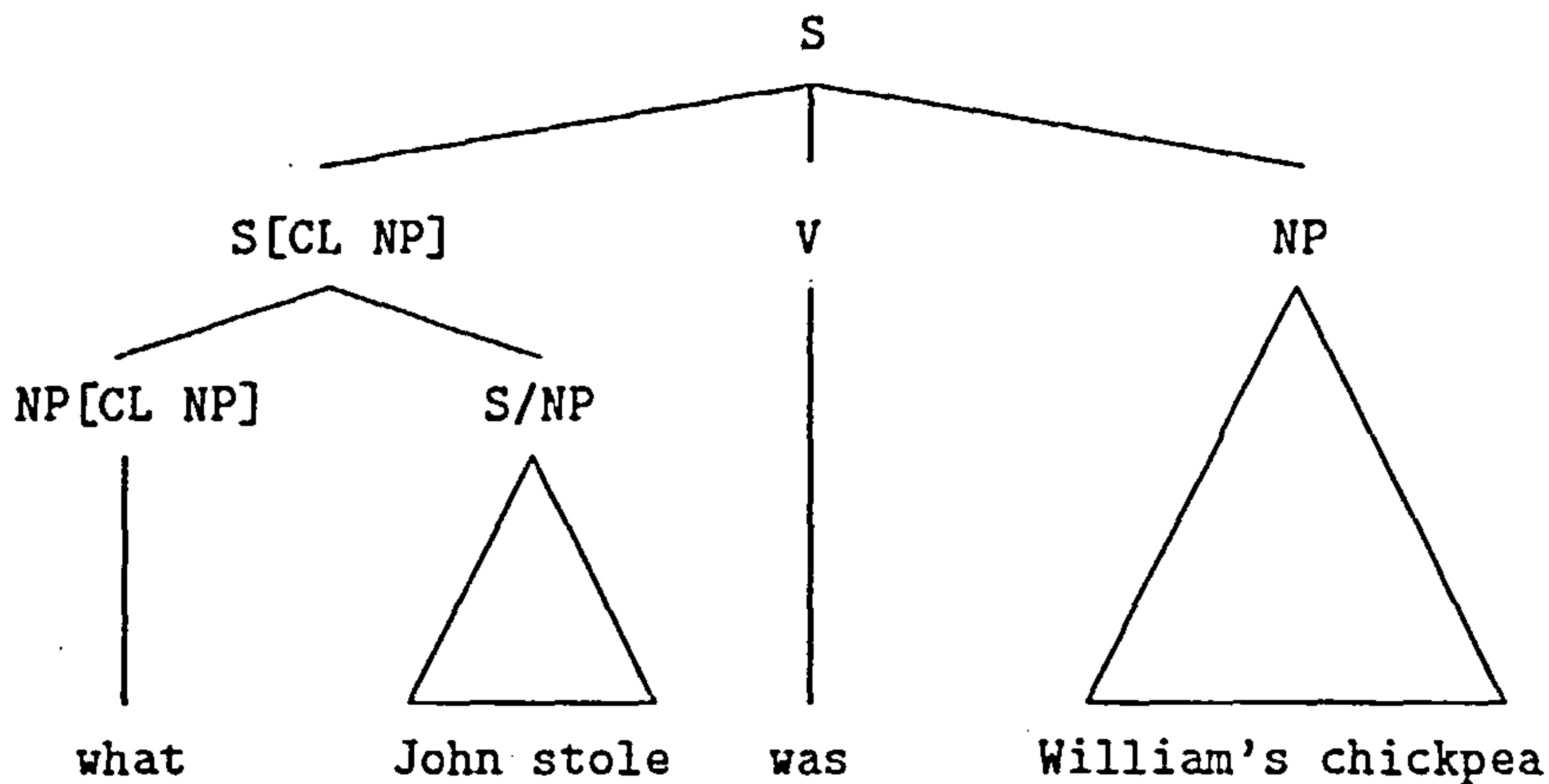


Figure 3.3: Phrase-marker for *wh*-cleft generated by rule (3-55)

The syntactic analysis we suggest for the *wh*-cleft *what John stole was William's chickpea* on the basis of rule (3-55) above appears in figure 3.3.

The first thing to note about the construction is that WHC has not been analysed as NP, as is sometimes done. The analysis of WHC as a sentential constituent is supported by the fact that the constituent cannot undergo movement such as subject-auxiliary inversion or subject-raising, both of which require the subject in question to be NP. (3-91a) is an example of a *wh*-cleft, while (3-91b) and (3-91c) are examples of the corresponding inverted and subject-raised sentences, respectively:

- (3-91) a What John is is a dancer.
 b *Is what John is a dancer?
 c *I believe what John is to be a dancer.

Halvorsen [1978:10] observes that 'vPs that both select a specificational [*wh*-cleft] reading cannot be conjoined', while vPs of *wh*-relative constructions can be. His examples are as follows:

- (3-92) a What John is doing is dangerous and is damaging to his health.
 b *What John is doing is buying tickets and is selling cars.

Halvorsen motivates this distinction on semantic grounds, arguing that examples such as (3-92b) are unacceptable because 'is is repeated in both conjuncts, and conjoined vPs with identical main verbs are in general ungrammatical' [Halvorsen 1978:11]. A

simpler explanation, however, falls out of our analysis of *wh*-clefts; as we noted above, *is buying tickets* and *is selling cars* are not VPs: they are not even constituents. This renders the ungrammaticality of (3-92) explicable in terms of the traditional constraints on constituent conjunction.

This analysis also explains behaviour noted by Higgins [1979] with respect to VP ellipsis: the copula of the cleft cannot be elided in conjoined constructions. In simple NP VP conjoined constructions, as in (3-93a), the main *v* of the second conjunct can be deleted under identity. In cases of conjoined *wh*-clefts such as (3-93b), however, the ellipsis is not possible, since no VP is present:

- (3-93) a John is stupid and Bill proud.
b *What John is is stupid and what Bill is proud.

Finally, the analysis of the copula itself as *v*, rather than VP, seems to be justified on the grounds that a non-cleft reading results in examples when the copula is modified in a way that suggests that it is not acting as an equative verb. Loss of the equative reading causes the cleft to be read as a *wh*-relative. Higgins notes that this may occur when a modal auxiliary precedes an embedded copula (as in (3-94a)), that copula is negated, as in (3-94b), or the copula is modified by a sentential adverb, as in (3-94c):

- (3-94) a What John is may be important.
b What John is isn't important.
c What John is is probably important.

Despite the attractions of using a flat structure for *wh*-clefts, we should point out that it is incompatible with the treatment of English word order proposed in Gazdar *et al.* [1985]. This is not a problem we will deal with here, but a solution is clearly required.

Having explained the motivation behind the analysis of the *wh*-cleft, we can turn to the business of differentiating the *wh*-cleft from other kinds of construction with which it has been conflated.

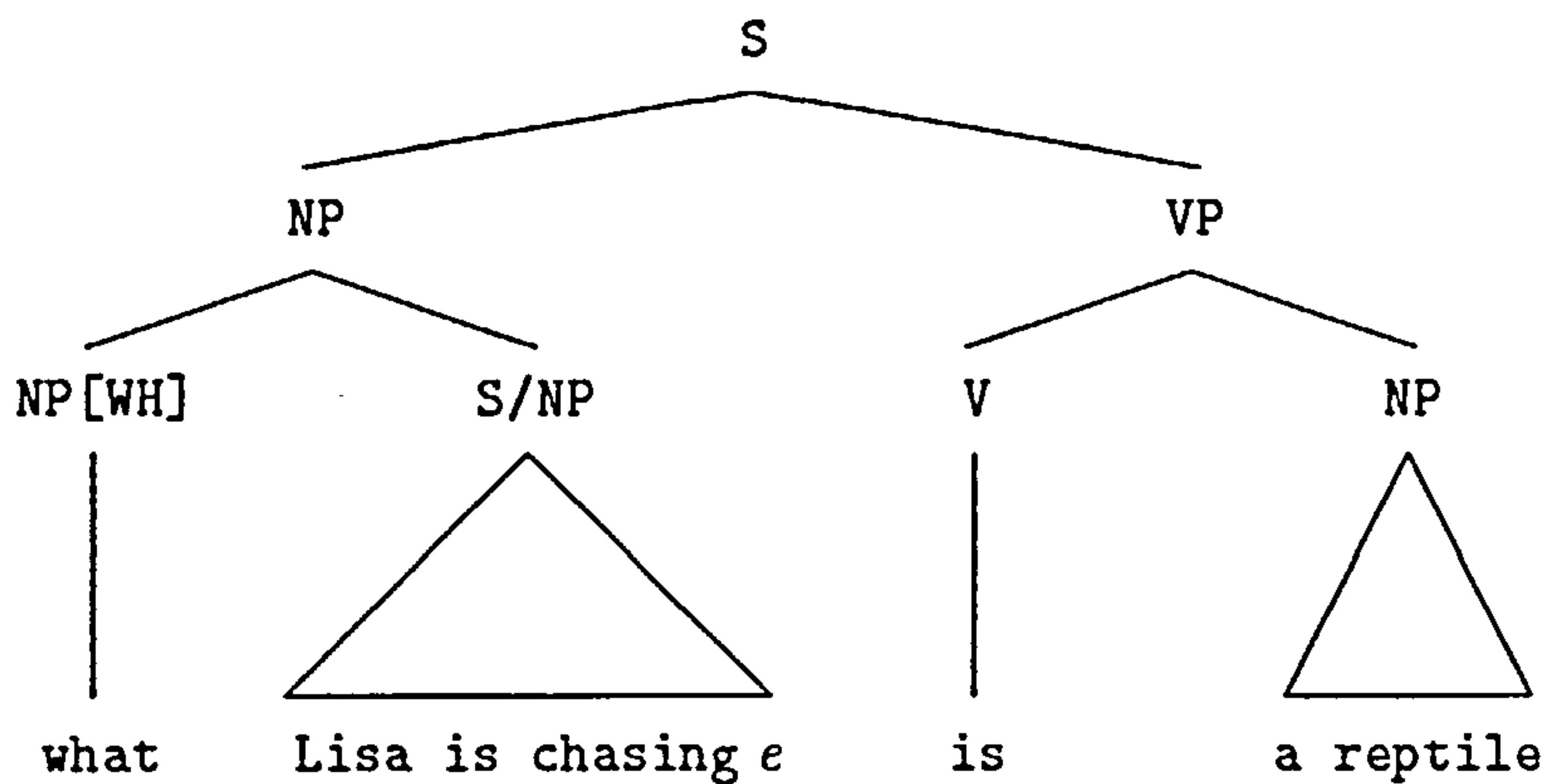


Figure 3.4: Phrase-marker for *wh*-relative construction

3.2.4 *Wh*-clefts and other *Wh*-Constructions

It is important to point out the syntactic differences between the *wh*-cleft and the *wh*-relative construction that gives rise to the predicational reading we described above. To recap briefly, a predicational reading results if the *wh*-clause of the *wh*-cleft is interpreted as a referring expression, and the complement of the copula is construed as an expression that predicates some property of the entity referred to. Sentences such as (3-95) have unambiguously predicational (*wh*-relative) readings; others, such as (3-96), are ambiguous between the *wh*-relative reading and the *wh*-cleft reading:

(3-95) What John likes is interesting.

(3-96) What John is is interesting.

The analysis we suggest for *wh*-relative constructions appears in figure 3.4, and can be motivated on the basis of many of the diagnostics we used as evidence for our analysis of *wh*-clefts. The example analysed in the figure is the *wh*-relative *what Lisa is chasing is a reptile*.

First of all, we argued that certain types of movement, such as subject-auxiliary inversion, and subject-raising, could not take place in *wh*-clefts because the WHC could not be analysed as NP. Our analysis of the *wh*-relative construction, however, where WHREL is NP, predicts that these kinds of movement will be acceptable. That this is the case is

demonstrated in (3-97) and (3-98). For clarity, we base these examples on the content of the unambiguous *wh*-relative in (3-95):

(3-97) Is what John likes important?

(3-98) I believe what John likes to be important.

We also argued against an analysis of the copula and head of the cleft as a VP constituent, demonstrating in support of this that operations such as VP conjunction could not be performed. Our analysis of the entire complement of the *wh*-relative as VP predicts that such conjunction should be possible in *wh*-relative sentences. As Halvorsen [1978] notes, this is indeed the case:

(3-99) What John likes is dangerous and is damaging to his health.

Further evidence for the VP analysis in the case of the *wh*-relative construction comes from Higgins, who notes that VP ellipsis can take place in a predication sentence, such as (3-100). As we saw above, this is not possible in the *wh*-cleft:

(3-100) What Mary likes is important and what John likes trivial in comparison.

(3-101) *What John is is stupid and what Bill is proud.

Finally, we noted Higgins' observation to the effect that *wh*-cleft modification of *wh*-cleft copulas could result in the loss of the *wh*-cleft reading. We captured this in the *wh*-cleft rule by describing the copula as an isolated *v* constituent. The VP analysis of the *wh*-relative copula and complement is supported by the fact that comparable examples of *wh*-relative modification are commonplace, and result in only the predictable changes of meaning:

- (3-102) a What John likes may be important.
b What John likes isn't important.
c What John likes is probably important.

We noted earlier that the *wh*-clause of the *wh*-cleft, here termed WHC, has often been confused with the *wh*-relative clause (WHREL) and the *wh*-interrogative (WHI). In this final part of the section on *wh*-clefts, we will turn our attention to a brief but conclusive

analysis that demonstrates that WHC can be analysed as neither WHREL nor WHI.⁶

There are three main parameters along which the clause types differ, as follows:

- The initial *wh*-words required by WHC, WHREL and WHI are distinct.
- Placement of prepositions in WHC and WHREL differs from that in WHI.
- Agreement in WHC patterns differently from that of WHREL and WHI.

We will address each of these areas in turn.

Distinctions in Initial *Wh*-Words

Accounts of *wh*-cleft syntax differ in their predictions regarding the kinds of *wh*-element that are acceptable as the initial item in the *wh*-cleft construction. Of the *wh*-elements that have been put forward as candidates, namely *who*, *what*, *where*, *when*, *why*, *how*, and *which*, Akmajian [1970] accepts all except *which*; Gundel [1977] includes *when* and *where* as well as *what*, while Higgins [1979] considers only *wh*-clefts with initial *what* to be grammatical. Of this set of possible *wh*-elements, only *which* appears to be positively ungrammatical:

(3-103) *Which I wanted was the green-coloured cut-out one.

(3-104) Who's in charge is Neil Kinnock.

(3-105) What they meant, of course, was that Conroy did not belong to the school that bash on regardless and hope for the best.

(3-106) Where Cheryl is going to wait is by the squash courts.

(3-107) When I was thinking of leaving was at about four.

(3-108) Why we're unhappy about this scheme is that we don't believe it will work.

(3-109) How I got to the tree was simply by adding one on each time.

⁶We do not intend to provide an exhaustive analysis of these clause types, since several authoritative works exist on the subject. See, for example, Bresnan and Grimshaw [1978].

Higgins [1979] notes that the suffix *-ever* cannot appear on the *wh*-element of WHC. For example:

(3-110) *Whatever John likes is sauerkraut.

Although WHC can (and usually does) begin with *what*, which is permitted both by the both WHREL and WHI, WHC differs from WHI in that WHC cannot begin with *whom*, *whose* or *which*:

(3-111) *Whom I am looking for is the manager.

(3-112) *Whose that car is is mine.

(3-113) *Which I asked for was the tinned one.

Quirk *et al.* [1985] claim that the following kinds of *wh*-words are acceptable initially in WHREL: *what*, *whatever*, *whichever*, *wherever*. In addition, *where*, *who*, *whoever*, *when*, *whenever*, *how*, *however*, *why*, *whom* and *whomever* seem to be possible, while *whyever* is unacceptable, as it is in all except WHI. Finally, *whose* seems to be marginally acceptable, and completely unacceptable with the suffix *-ever*. These cases can be seen in the following examples:

- (3-114) a I'll do what/whatever you do.
b I'll choose *which/whichever you choose.
c I'll go where/wherever you go.
d Just ask who/whoever you know to come.
e Come when/whenever you're ready.
f Do it how/however you can.
g He did it why/*whyever you did it.
h I'll talk to whom/whomever you are to see.
i ?I'll buy whose/*whosever I want.

WHI, on the other hand, can begin with *what*, *which*, *where*, *who*, *when*, *how*, *why*, and *whose*. All of these except *whom* and *whose* can take the suffix *ever*. *Whose* does not in general seem to be acceptable with *-ever* even outside the context of the cleft. The unacceptability of *whomever*, though, appears to be a constraint particular to WHI:

- (3-115) a What/Whatever is she doing?
 b Which/Whichever will he choose?
 c Where/Wherever will she go?
 d Who/Whoever is going to come?
 e When/Whenever will they be ready?
 f How/However will she manage?
 g Why/Whyever did they do it?
 h Whom/*Whomever are you going to see?
 i Whose/*Whosever is that?

Finally, only in WHI can *whose*, *which* and *what* appear with NP. In examples (3-116)–(3-118), the WHI appears in the (a) examples, WHREL in the (b) examples, and WHC in the (c) examples:

- (3-116) a I wonder whose cat that is.
 b *I'll love whose cat you love.
 c *Whose cat John wanted was Bill's.
- (3-117) a I wonder which cat that is.
 b *I'll love which cat you love.
 c *Which cat John wanted was the grey one.
- (3-118) a I wonder what cat that is.
 b *I'll love what cat you love.
 c What cat John wanted was the grey one.

Therefore, although the *wh*-word *what* is acceptable in all three classes of clause, the sets of acceptable *wh*-initiators are otherwise distinct to a large extent. A table showing the differing ranges of acceptable *wh*-elements of WHREL, WHC and WHI appears in figure 3.5. In this table, unacceptable *wh*-elements are indicated by asterisks, while acceptable ones are ticked.

Differences in Placement of Prepositions

Quirk *et al.* [1985:1060] note differences in the acceptability of prepositions in certain positions in WHI and WHREL. Prepositions occurring in WHREL have to appear after the *wh*-word, as is shown in (3-119):

<i>wh</i> -word	WHI	WHR	WHC
<i>who</i>	✓	✓	✓
<i>whoever</i>	✓	✓	*
<i>whose</i> + NP	✓	*	*
<i>what</i>	✓	✓	✓
<i>whatever</i>	✓	✓	*
<i>what</i> + NP	✓	*	*
<i>when</i>	✓	✓	✓
<i>whenever</i>	✓	✓	*
<i>where</i>	✓	✓	✓
<i>wherever</i>	✓	✓	*
<i>how</i>	✓	✓	✓
<i>however</i>	✓	✓	*
<i>why</i>	✓	✓	✓
<i>whyever</i>	✓	*	*
<i>which</i>	✓	*	*
<i>whichever</i>	✓	✓	*
<i>whose</i>	✓	?	*
<i>whosever</i>	*	?	*
<i>whom</i>	✓	✓	*
<i>whomever</i>	*	✓	*
<i>which</i> + NP	✓	*	*

Figure 3.5: Comparison of acceptable *wh*-elements in WHREL, WHC and WHI

- (3-119) a They ate what they paid for.
 b *They ate for what they paid.

WHI, on the other hand, allows prepositions to be fronted, as in (3-120):

- (3-120) a I don't know who they lent the money to.
 b I don't know to whom they lent the money.

In this respect, WHC and WHI display different behaviour, WHC in fact approximating more closely the pattern of WHREL: like WHREL, WHC doesn't tolerate the appearance of the preposition prior to the *wh*-word. This is exemplified in (3-121):

- (3-121) a What he was driven to was suicide.
 b *To what he was driven was suicide.

In case it appears that the restriction is on prepositions such as this appearing sentence-initially, the WHI example in (3-122)⁷ allows prepositions to be fronted:

- (3-122) To what he owes allegiance is a mystery.

It seems that the unacceptability of these examples is due to the fact that the clause from which the preposition is extracted is headless. A clefted copular sentence with *the one* as head, such as (3-123), does permit this extraction:

- (3-123) a The one I was complaining to was him.
 b The one to whom I was complaining was him.

In this respect, therefore, WHC resembles WHREL more closely than WHI.

Differences in Agreement

Halvorsen [1978], in his discussion of Faraci's [1971] claim that the WHC is in fact interrogative, argues that both embedded interrogatives and *wh*-clefts always have singular marking on the copula, while copulas in *wh*-relative constructions can have either singular or plural marking. First of all, it is quite clear that Halvorsen's claim regarding obligatory singular marking on *wh*-cleft copulas is incorrect. *Wh*-clefts permit plural

⁷I am grateful to Graeme Ritchie for suggesting this example to me.

marking too, as shown in (3-124), apparently indicating agreement with either the *wh*-clause or with the cardinality of the head element:

(3-124) What we were looking at was/were some old photos of Mum.

Secondly, the interrogative, too, may condition plural marking on a succeeding copula if two interrogatives are conjoined. This suggests that agreement is not with the cardinality of the NP within the *wh*-clause, but with the *wh*-clause itself—that is, with the number of questions that are being asked in the sentence:

- (3-125) a Which boys are ahead in the race is/*are unclear.
b Which boys are ahead in the race and what time it's going to end are both unclear.

As Halvorsen shows, however, plural agreement is obligatory in *wh*-relative constructions, if a plural NP appears in the *wh*-clause complement. This indicates that the cardinality of that complement controls agreement:

(3-126) What John just stepped on *is/are living beings.

In fact, the agreement behaviour of the *wh*-cleft is identical to neither the pattern shown by WHI, nor that of WHREL. As we saw above, when a plural element appears in the complement of a *wh*-cleft, plural agreement is optional. This behaviour, shown in (3-127), is in marked contrast to that of WHREL, demonstrated in (3-126):

(3-127) What John just stepped on was/were living beings.

Number agreement behaviour in *wh*-clefts is therefore different from that occurring in either WHREL or WHI clauses.

In summary, then, we have observed several differences between WHC and two other kinds of clause with which it has been equated, namely WHREL and WHI. Our conclusions were as follows:

- WHC, WHREL and WHI take overlapping but different sets of *wh*-elements.
- Prepositions may appear prior to the *wh*-element in WHI, but not in WHC and WHREL.

- Agreement in WHREL is with the complement of the copula, in WHC with either that element or with the *wh*-clause, and in WHI with the number of *wh*-clauses that appear.

We conclude, therefore, that an analysis of WHC in terms of either WHREL or WHI is not supported by the data. In fact, the *wh*-cleft *wh*-clause seems to occupy a unique position in the range of *wh*-clauses, in the following way. It seems plausible to posit a continuum—on an intuitive semantic level at least—from interrogative, through *wh*-cleft, to free relative, in terms of the referential behaviour of the three kinds of clause. (3-128) shows each type respectively:

(3-128) I asked what he wanted.

(3-129) He said that what he wanted was a bag of crisps.

(3-130) I gave him what he wanted.

It seems that what differentiates between the three examples—which all belong to different classes of clause, as far as we have been able to establish—is that WHI is asking for a referent, the *wh*-cleft sentence structure assigns a referent to WHC, while WHREL is being used as a referring expression. Referentiality—or at least capability of carrying the features that pick out a referent—seem to be behind part of the distinction between the predication and *wh*-cleft sentences: number agreement and ability to take the suffix *-ever* at least.

3.2.5 Agreement in *Wh*-clefts

Although we touched briefly on matters of agreement in the discussion of *wh*-cleft *wh*-clauses above, some aspects of agreement in *wh*-clefts remain to be summarised.

A clue to the syntax of *wh*-clefts lies in how agreement of the copula in person, number and tense is controlled. Sornicola [1988:348] claims that agreement is controlled by the head of the *wh*-clause in cases where items like *the one/the ones* are present, or by the relative pronoun itself (the *wh*-word) when there is no nominal head in the *wh*-clause. Sornicola's claim certainly seems to be true of *wh*-clauses that have heads, since person

and number agrees with that of the head in each case:

(3-131) The ones I want *is/are the blue ones.

(3-132) The one I want is/*are the blue one.

The claim that the *wh*-word in the headless WHREL clause can also control agreement, however, is more problematic. First of all, the *wh*-clause is morphologically unmarked for number. It may carry number agreement information lexically, on the other hand, but this is almost impossible to evaluate.

Gundel [1977] takes the view that the *wh*-cleft copula agrees with the head of the *wh*-cleft. This appears to be the case in (3-133), where there is no number information morphologically marked on the *wh*-clause, and none is available from any semantic antecedent. Plural agreement, however, still occurs on the copula:

(3-133) What we were looking at were some old pictures of Mum.

In cases where the cleft head is plural, singular marking of the copula (such as in (3-134)) is also possible:

(3-134) What we were looking at was some old pictures of Mum.

On the basis of (3-134), it seems that the one option for agreement is with the *wh*-clause, which, as it is not marked for number, will result in a neutral singular. To accommodate examples such as (3-133), we can add the second option of agreement with the cleft head. The way this agreement alternates, and the features that control it, are as yet poorly understood.

However, it seems that things are more complex than simple alternation of agreement control between head and *wh*-clause. There are also examples where agreement on the copula doesn't seem to occur at all, in spite of the relevant information being available. This is the case in (3-135a). (3-135b), which does display agreement, is also acceptable:

- (3-135) a Who those people are is Mike's friends from Sheffield.
b Who those people are are Mike's friends from Sheffield.

This lends weight to an analysis such as that suggested for the *wh*-interrogative, in

which agreement appears to be with the physical number of *wh*-clauses that appear, regardless of their cardinality. We could say for the *wh*-cleft that, when the copula does not agree with the head, it agrees with fact that there is only a single *wh*-clause.

Agreement in tense is more clear-cut than agreement in number. In *wh*-clefts, the copula is required to agree in tense with the *wh*-clause. In (3-136a) and (3-136b), this 'tense harmony' (cf. Higgins [1979]) is observed:

- (3-136) a What John has got is a tree.
b What John had was a tree.

In (3-138), however, the copula and the *wh*-clause do not agree in tense, and as Higgins [1979] predicts, a *wh*-relative reading results:

- (3-137) a What John has got will be a tree.
b What John will get has been a tree.

3.2.6 Summary

In this section, a rule for *wh*-clefts was suggested, and its predictions explored. It was concluded that a superficially similar construction to the *wh*-cleft, the *wh*-relative construction, required a syntactic analysis distinct from that of the *wh*-cleft, and the distinction was motivated on the basis of several diagnostics for constituency. The *wh*-clause of the *wh*-cleft was examined in some detail, and differentiated from two other clause types with which it has been identified, the *wh*-relative and the *wh*-interrogative clause, on the basis of three factors: distinctions in the initial *wh*-elements available for each, distinctions in terms of the placement of prepositions, and differences in agreement behaviour.

As in the case of the *it*-cleft, it was proposed that the head of the *wh*-cleft could be analysed as XP, and we examined the constraints on what constituent types could instantiate XP. The range of possible constituent types was in general broader than that of the *it*-cleft, although the following were ruled out, some of them on the basis that they induced the *wh*-relative reading for the *wh*-cleft: AP when the main verb of the *wh*-clause was not an intensive verb, non-volitional manner adverbials, finite VP, and all anaphoric items.

3.3 The Reverse *Wh*-Cleft

We turn now to the least studied of the three kinds of cleft construction that are the subject of this thesis: the reverse *wh*-cleft (also termed INVERTED PSEUDO-CLEFT). An example of this cleft type appears in (3-138):

(3-138) This is what annoys me most.

As in the discussion of the *wh*-cleft, the relative-clause-like part of this construction will be termed the WH-CLAUSE, while the pre-copular part will be referred to as the cleft's HEAD.

3.3.1 Phrase-Structure Rules for Reverse *Wh*-clefts

For reverse *wh*-clefts, we suggest the following rules:

(3-139) $S \longrightarrow XP\ VP[AGR\ XP]$

(3-140) $VP[AGR\ XP] \longrightarrow V\ S[CL\ XP]$

These rules give rise to a phrase-marker such as that in figure 3.6, which analyses the example *that is what I want*.

The motivation behind this analysis is as follows. There are important similarities between this analysis and that of the *wh*-cleft, presented above. As in the case of the *wh*-cleft, there are two dependencies to be taken care of. The first is a straightforward syntactic dependency between the *wh*-filler in the *wh*-clause and the corresponding gap. The second requires the agreement of the XP head of the cleft with the XP value in the *wh*-clause. Since the XP constituent in (3-139) controls the VP, the Control Agreement Principle [Gazdar *et al.*1985] will ensure that any further instantiation of XP will be identical to the eventual specification of AGR on VP.

We have shown in the case of the *wh*-cleft that if the specification [CL XP] appears on the mother category, it will ensure a *wh*-filler of category XP. To take care of the corresponding match between the value of [AGR] on the VP node and the value of the

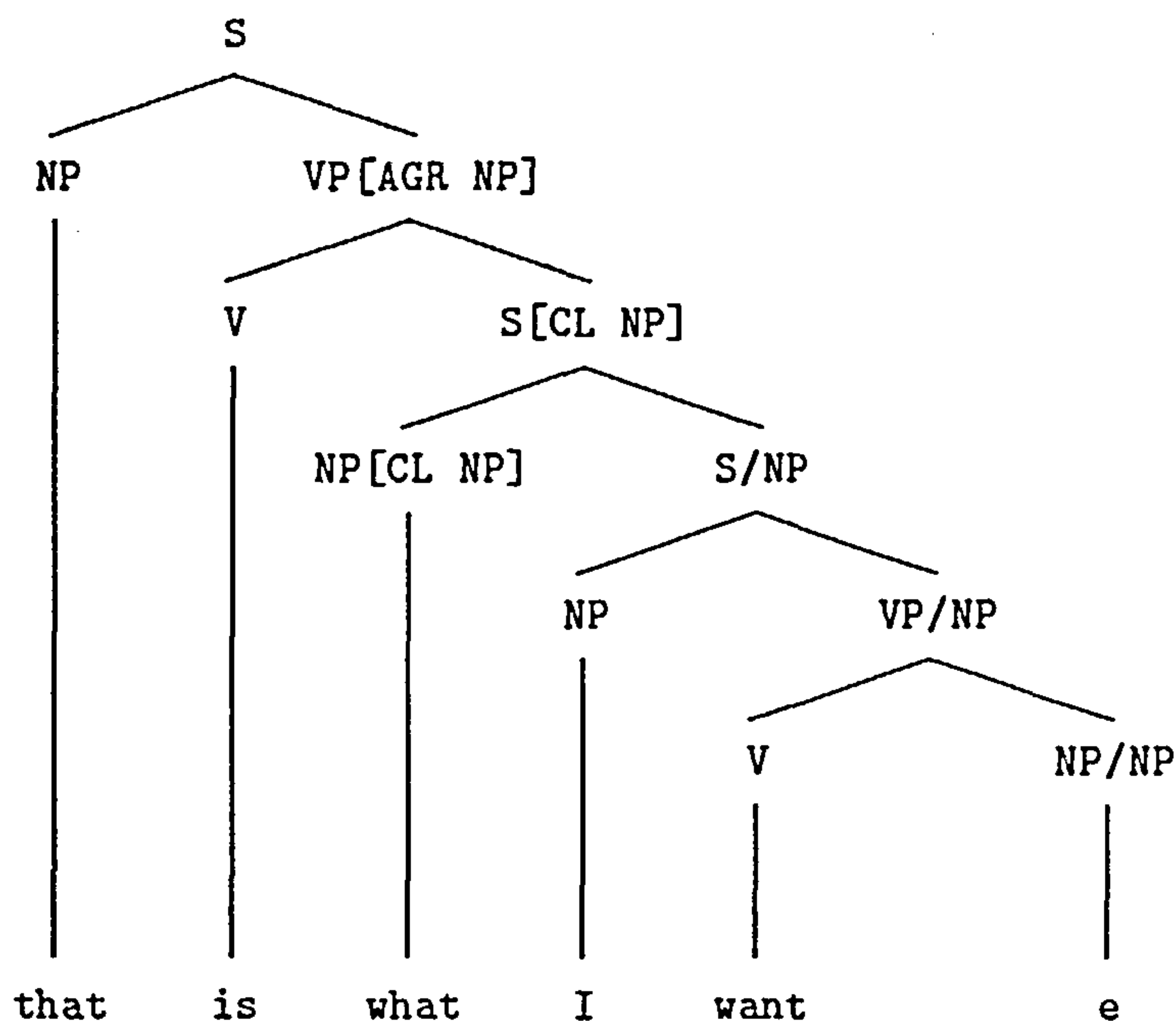


Figure 3.6: Phrase-marker for reverse *wh*-cleft

feature [CL] on the daughter of that VP, we will assume that this too can be handled by the appropriate FCR. How this is worked out in detail is not something we will investigate here, however.

An important distinction exists between the analysis for the reverse *wh*-cleft proposed here and that of the *wh*-cleft, however: the reverse *wh*-cleft copula and its complement are analysed as VP. Evidence for this analysis is that, while in the case of the *wh*-cleft the copula and its complement cannot be conjoined as VP, in the reverse *wh*-cleft it is possible to do this. Examples of *wh*-cleft and reverse *wh*-cleft respectively appear in (3-141):

- (3-141) a *What John is doing is buying tickets and is selling cars.
 b This hat is what I wanted and is what I still want.

Further support for the simple XP VP analysis comes from subject-auxiliary inversion and subject raising of appropriate examples. As we noted in previous sections, for either kind of movement to take place the copula and its complement have to form a VP, while the initial constituent of the cleft has to be a sister of that VP acting as subject. Both kinds of movement are possible in reverse *wh*-clefts. In (3-142), (a) is a canonical reverse

wh-cleft, (b) is the corresponding sentence with subject-auxiliary inversion, and (c) is a subject-raised version:

- (3-142) a This is what you wanted.
b Is this what you wanted?
c I believe this to be what you wanted.

In the discussion of *wh*-clefts, we observed that copula deletion in conjoined structures was not possible, since the copula was not part of VP:

- (3-143) *What John is stupid and what Bill is proud.

Our analysis predicts that copula deletion in the second conjunct will be acceptable in reverse *wh*-clefts, since the copula is an ordinary *v* within VP. This turns out to be the case:

- (3-144) That is what I wanted and this what I got.

A further point in support of the VP analysis for the copular constituent is the fact that the copula can, unlike that of the *wh*-cleft, take the kinds of modification that are expected in ordinary VPs. Examples are as follows: the starred examples, while grammatical, are unambiguously *wh*-relative readings of the sentences involved:

- (3-145) a *What John is doing is maybe working.
b Working is maybe what John is doing.
c *What John is doing isn't working.
d Working isn't what John is doing.
e *What John is doing is probably working.
f Working is probably what John is doing.

Finally, the agreement behaviour of the reverse *wh*-cleft suggests that the XP VP analysis is the correct one, since it is simply the case that, as in canonical declaratives, the copula agrees with the subject in all cases:

- (3-146) Some photos of Mum *was/were what we were looking at.

- (3-147) A photo of Mum was/*were what we were looking at.

As a final point before we turn to the examination of reverse *wh*-cleft heads, it should be noted that the *wh*-clause of the reverse *wh*-cleft (RWHC, for simplicity) is not comparable to either the *wh*-relative (WHREL) nor the *wh*-cleft *wh*-clause (WHC) in terms of the initial *wh*-elements possible. RHWC behaves like WHC and unlike WHREL in that both *whom* and the suffix *-ever* are unacceptable:

- (3-148) a *That's whom I'm looking for.
b *That's whatever I want.

Unlike either of the other two clause types, however, *whose* with NP seems to be acceptable:

- (3-149) That's whose car I borrowed.

Having examined the syntactic structure of the reverse *wh*-cleft, we can look more closely at the constituents that can appear as XP: the cleft head.

3.3.2 The Instantiation of XP

Looking first at the grammaticality of AP as cleft head, we find the case of the reverse *wh*-cleft similar to that of the *it*-cleft. While AP is acceptable as head in many *wh*-clefts, in reverse *wh*-clefts AP is often marginal or unacceptable:

- (3-150) a What your pie was was absolutely delicious.
b ?It was absolutely delicious that your pie was.
c ?Absolutely delicious was what your pie was.

- (3-151) a What he became was morose.
b *It was morose that he became.
c ?Morose was what he became.

- (3-152) a How it felt was furry.
b *It was furry that it felt.
c ?Furry was how it felt.

In cases where AP is marginal in the *wh*-cleft, it is unacceptable in the corresponding reverse *wh*-cleft:

(3-153) a ?What John seems is unhappy.

b *Unhappy is what John seems.

(3-154) a ?What John remained was morose.

b *Morose was what John remained.

ADVP as head of the reverse *wh*-cleft is subject to similar constraints to those that apply in the case of the *wh*-cleft. As we saw earlier, only volitional adverbials that fall into Quirk *et al.*'s [1985:482] categories of MEANS and INSTRUMENT adverbials, and that can therefore be bound by *how*, can form the head of a reverse *wh*-cleft. Means adverbials refer to the means by which an action is done, including phrases such as *by bus* and *through insight*. Instrument adverbials include phrases such as *with a fork* and *using a dictionary*. Examples of reverse *wh*-clefts with means adverbials as head appear in (3-155), while the clefts in (3-156) have instrument adverbials as head:

(3-155) a By bus was how we got here.

b Through insight was how he reached a solution.

(3-156) a With a fork is how you turn over the topsoil at first.

b Using a dictionary was how she solved most of the clues.

At first glance, *-ly* adverbs seem to be marginal or unacceptable in general:

(3-157) ?Carefully was how he rowed.

(3-158) *Luckily was how he won.

However, as we discussed in the case of the *wh*-cleft, the appropriate generalisation is that *-ly* adverbs fall Quirk *et al.*'s non-volitional class of MANNER adverbials, and cannot therefore be bound by *how*. The constraint also applies to PP as manner adverbial:

(3-159) ?With great aplomb was how he got there.

The constraints on ADVP, therefore, are equivalent to those of the *wh*-cleft.

In the case of the *wh*-cleft, we saw that sentential complements of both factive verbs such as *regret* and those of non-factive verbs such as *write* made acceptable heads. Reverse *wh*-clefts behave more like *it*-clefts in this respect, as in the *it*-cleft the complements of non-factives are unacceptable, and in reverse *wh*-clefts they are marginal:

(3-160) ?That he was coming home was what John wrote.

(3-161) That he was coming home was what John regretted.

Like the case of *it*-clefts, negative polarity items make bad reverse *wh*-cleft heads. As we saw above, these are acceptable in *wh*-cleft:

- (3-162) a *It is any marbles that John doesn't want.
b *Any marbles are what John doesn't want
c What John doesn't want are any marbles.

Quantifier phrases are similarly unacceptable:

(3-163) *Every film of Coppola's is what I've seen.

Finally in this discussion of head types, the behaviour of reverse *wh*-clefts with respect to non-finite VPs is as follows. As in the case of *it*-clefts and *wh*-clefts, purposive *to*-infinitives, *-ing* participle phrases as subject, and passive participle phrases are all acceptable. Examples of each respectively appear in (3-164):

- (3-164) a To get shaving cream was what he went out for.
b Making all those beds was what was the worst.
c Jammed under the settee was where I found your wallet.

The reverse *wh*-cleft and the *wh*-cleft can both take non-purposive *to*-infinitives, while we saw in section 1 that this was not possible for *it*-clefts:

- (3-165) a What I'd prefer is to get a more appley green.
b To get a more appley green is what I'd prefer.
c *It's to get a more appley green that I'd prefer.

Bare infinitives, on the other hand, while acceptable in *wh*-clefts, are only marginally acceptable in reverse *wh*-clefts:

- (3-166) a What she'd like to do is leave straight away.
b ?Leave straight away is what she'd like to do

Finally, while VPs with progressive aspect seem to be as acceptable in *wh*-clefts as non-progressive VPs, only the former seem to be acceptable in reverse *wh*-clefts. (3-167) shows the progressive examples, (3-168) the non-progressive:

- (3-167) a What John is doing is making the beds.
 b Making the beds is what John is doing.
- (3-168) a What Mary does is see a psychiatrist.
 b ?See a psychiatrist is what Mary does.

The range of heads possible for the reverse *wh*-cleft might seem on the evidence above to be more restricted than that of the *wh*-cleft. However, the anaphoric heads available to the reverse *wh*-cleft render this construction rather more flexible than it first appears. As we show in the tables given in section 3.5, the vast majority of reverse *wh*-cleft heads are in fact composed of the pro-forms *this* and *that*, which, as we noted in section 3.2, are not available to the *wh*-cleft. In practice, while the constraints on the appearance of certain kinds of constituent as head still operate, the same content can often be specified, if suitable discourse conditions exist, by *this* or *that*. For example, (3-169) specifies the content of a bare infinitive, which we saw above to be an unacceptable reverse *wh*-cleft head in itself:

- (3-169) a *Leave straight away was what she did.
 b What she wanted to do was leave straight away, so that's what she did.

(3-170) shows a similar treatment of the content of an *-ly* adverb:

- (3-170) a *Energetically was how he ran.
 b I thought he would run energetically, and that's how he ran.

Although the complements of non-factive verbs were themselves unacceptable as reverse *wh*-cleft heads, *that* can be used to specify the same content:

- (3-171) I told them he'd write that he was coming home, and that's what he wrote.

Anaphoric respecification can be used to create a new grammatical role for the specified information. For example, (3-172) respecifies previously occurring content and uses it adjunctively within the reverse *wh*-cleft, even when the original specification of the information would not have been acceptable as an adjunct (cf. (3-173)):

- (3-172) The cat came in and the dog went out. That was why I couldn't shut the kitchen door.
- (3-173) *The cat came in and the dog went out was why I couldn't shut the kitchen door.

Note that the reverse *wh*-cleft can take pronominal *it* and *which* as cleft head, which, as we saw in the previous sections, neither of the other two cleft types can:

- (3-174) a He announced—and it is what I cannot forgive him for—that he had spiked Annabel’s Perrier.
 b *He announced—and it is it that I cannot forgive him for—that he had spiked Annabel’s Perrier.
 c *He announced—what I cannot forgive him for is it—that he had spiked Annabel’s Perrier.
- (3-175) a The most frightening part, which of course is what concerns me, is that nobody knows this apart from Deeping.
 b *The most frightening part, it is which of course that concerns me, is that nobody knows this apart from Deeping.
 c *The most frightening part, what of course concerns me is which, is that nobody knows this apart from Deeping.

This constraint is related to consideration of what constitutes acceptable serial ordering of information. We address this issue in detail in chapter 5.

Finally, note that uniquely among the clefts the reverse *wh*-cleft is unable to accept negated head constituents of any kind:

- (3-176) a * Not John was who did it
 b *Not only the wind was what made the noise
 c *Not until three was when she came

The distinctions between all three types of cleft in terms of what can instantiate the head XP constituent will be summarised at the end of the chapter.

3.3.3 Summary

The analysis given for the reverse *wh*-cleft has been the simplest of all, namely XP VP. As far as the instantiation of XP is concerned, we found the range of constituents possible closer to that of the *it*-cleft than to the *wh*-cleft. As well as the usual prohibitions on constituents that were not major phrasal types, the following constituent types were unacceptable as XP: AP, non-volitional manner ADV and ADVP, negative polarity NP, quantifier phrases, non-finite VPs with non-progressive aspect, finite VPs, and bare infinitives. Sentential complements of non-factive verbs were marginal as XP. As in all

three cleft types, sentential adverbs were not acceptable. The range of anaphoric elements that could appear as head of the reverse *wh*-cleft was the broadest of all three types of cleft, and included, uniquely among the clefts, pronominal *which*.

In section 3.5, a comparative summary of the syntax of the three cleft types is presented. First, however, we will take a brief look at the frequency of the main variants appearing in the corpus data.

3.4 A Note on the Corpus Data

In this section, the data taken from the joint corpora are presented, taken from a total of 245 *it*-clefts, 162 *wh*-clefts, and 160 reverse *wh*-clefts. Two frequency counts were carried out, the most important being a count of the frequencies of various cleft head types. The second count gives the frequency and range of *wh*-elements that could appear as relativisers or complementisers in the three types of cleft.

It might be expected that the examples from a corpus of naturally-occurring data would yield many strange and quasi-grammatical strings. In fact, the data found in the corpora were quite ordinary in this respect. The interesting points to note, however, are the narrowness of the range of data that appeared, compared to the broad range of possible grammatical strings that could have occurred, and the marked differences between the cleft types with respect to the appearance of particular syntactic variants.

Constituents appearing as Head

In the frequencies below we omit clefts that have more than one constituent, such as two conjoined NPs, as head. We distinguish between human and non-human NPs, and between various types of NP in the case of reverse *wh*-clefts. For each category, three figures are given: the frequency from the spoken corpus, the frequency from the written corpus, and the total frequency from the corpus as a whole.

The interesting points to note here are the distinctions between the three cleft types. All three cleft types, as might be expected, have a high proportion of NP heads. The

	spoken [91]	written [154]	total [245]
NP [+HUMAN]	34	39	74
NP [-HUMAN]	44	41	85
Total NP	78	80	158
PP	4	46	50
ADVP	9	27	36
S	0	1	1

Figure 3.7: Constituent types appearing as heads of *It*-clefts

	spoken [99]	written [63]	total [162]
NP [+HUMAN]	2	1	3
NP [-HUMAN]	38	30	68
Total NP	40	31	71
PP	0	0	0
ADVP	2	0	2
S	29	23	52
VP [-FIN]	29	9	38

Figure 3.8: Constituent types appearing as heads of *Wh*-clefts

reverse *wh*-cleft, however, in spite of the wide range of possible head constituents, has NP heads in all but one case. Notable in *it*-clefts is the very low frequency of s heads, and the complete non-appearance of non-finite VPs of any kind. *Wh*-clefts, on the other hand, have a very high proportion of s heads and non-finite VPs, but, surprisingly, no PPs at all. In none of the cleft types do *-ly* adverbs or APs appear as head.

	spoken [122]	written [38]	total [160]
<i>that</i>	89	13	102
<i>this</i>	27	16	43
other NP	5	7	12
<i>which</i>	0	2	2
PP	1	0	1

Figure 3.9: Constituent types appearing as heads of Reverse *Wh*-clefts

	spoken [91]	written [165]	total [256]
<i>that</i>	44	102	146
<i>who</i>	18	33	51
0	24	6	30
<i>when</i>	2	2	4
<i>which</i>	3	19	22
<i>why</i>	*		
<i>how</i>	*		
<i>where</i>	*		
<i>what</i>	*		
<i>which</i> + PREP	0	3	3

Figure 3.10: Constituent types appearing as *it*-cleft *wh*-elements

Relativisers/Complementisers Appearing

The total occurrences of various relativisers/complementisers found in the corpora appear below. The notation is that of the previous section. The sizes of the corpora are slightly different from those shown above since the entire corpus, including those clefts with conjoined heads, has been included in this count. *Wh*-elements that are ungrammatical for the cleft type in question are indicated by asterisks.

In spite of the wide range of *wh*-elements predicted to be possible for the *wh*-cleft, the range appearing is in fact very narrow: almost exclusively *what*, with a very few instances of *how*, *who* and *why*. The spread of reverse *wh*-clefts is more broad, but again with the vast majority of examples featuring *what*. For *it*-clefts, *that* is the most common, followed by *who* and complementiser deletion.

3.5 The Syntax of Cleft Constructions: Summary

In this chapter, we conducted an analysis of each of the three cleft types in turn, suggesting appropriate phrase-structure rules for the generation of each. In the case of the *it*-cleft, we adopted an existing pair of rules suggested by Gazdar *et al.* [1985], which capture in particular the fact that the complement of the *it*-cleft seems to behave like a relative clause when the cleft head is NP, but like a complementised sentence when the cleft head is any other category. For *wh*-clefts, we suggested an analysis in which

	spoken [99]	written [72]	total [171]
<i>that</i>	*		
<i>who</i>	1	0	1
0	*		
<i>when</i>	0	0	0
<i>which</i>	*		
<i>why</i>	2	0	2
<i>how</i>	1	0	1
<i>where</i>	0	0	0
<i>what</i>	95	72	167
<i>which</i> + PREP	*		

Figure 3.11: Constituent types appearing as *wh*-cleft wh-elements

	spoken [122]	written [38]	total [160]
<i>that</i>	*		
<i>who</i>	3	0	3
0	*		
<i>when</i>	2	0	2
<i>which</i>	0	0	0
<i>why</i>	14	6	20
<i>how</i>	7	3	10
<i>where</i>	13	1	14
<i>what</i>	83	28	111
<i>which</i> + PREP	*		

Figure 3.12: Constituent types appearing as reverse *wh*-cleft wh-elements

the *wh*-clause was treated as a special clause type, on the basis of its lack of similarity with other superficially similar types of clause. A syntactic dependency existed within the *wh*-clause between the *wh*-filler and a gap of the same category. In addition, a dependency existed between the *wh*-filler in the *wh*-clause and the head of the *wh*-cleft. We took care of this by specifying a category-valued feature [CL XP] on the *wh*-clause, inherited by the *wh*-filler, which matches the current value of XP on the cleft head. The analysis of the reverse *wh*-cleft was conducted in a similar way, except in this case agreement between the XP category of the *wh*-filler in the *wh*-clause, and the cleft head, was ensured by means of the specification AGR[XP] on the VP. This VP itself is worthy of note, since we argued that, unlike the case of the *wh*-cleft, the reverse *wh*-cleft copula and its complement could be analysed as a single VP constituent, rather than as V XP as in the case of the *wh*-cleft.

In each section, we devoted some time to examining what kinds of constituent could appear as XP—that is, as head of the cleft. We noted that the range of constituents it was possible to have as XP was different for each cleft type. Some of the reasons behind the distinctions may be pragmatic; others may be the result of syntactic constraints that remain for further research. The possible head constituents of each cleft type, however, may be important determinants of their relative functions in discourse, in terms of the choice of a cleft type to convey a particular message. The range of possible head types are compared in figure 3.13.

Having filled in the syntactic background to our study of clefts we can turn to perhaps the most powerful factor determining the function of these constructions in discourse: their semantics.

	<i>it</i> -clefts	<i>wh</i> -clefts	reverse <i>wh</i> -clefts
NP	✓	✓	✓
Finite VP	*	*	*
Purposive <i>to</i> -infinitive	✓	✓	✓
Non-purposive <i>to</i> -infinitive	*	✓	✓
- <i>ing</i> as VP complement	*	✓	✓
Subject <i>ing</i> participle	✓	✓	✓
Passive participle	✓	✓	✓
Bare infinitive	*	✓	?
NP + <i>not until</i>	✓	*	*
XP + <i>not only/just</i>	✓	✓	*
Universally-quantified NP	*	✓	*
Negative polarity NP	*	✓	*
AP	? *	✓ (most)	? *
- <i>ly</i> ADV	*	*	✓
ADVP	✓	some	some
Sentential ADV	*	*	*
s complement of factive v	✓	✓	✓
s complement of non-factive v	*	✓	?
PP	✓	✓	✓

Figure 3.13: Possible Head Constituents by Cleft Type

Chapter 4

Aspects of Cleft Semantics

In this chapter, we explore how the cleft itself, and its interaction with context, can be characterised in semantic terms. In what follows, we will argue for a view of clefts as differentiated from declaratives on the basis of several features that can be captured semantically. The first, and most crucial, is that the cleft is a PRESUPPOSITIONAL construction in a way that simple declaratives are not. We can characterise the content of the part of the cleft that conveys the presupposition, the *wh*-clause or cleft complement, quite simply. For example, for the three cleft constructions in (4-1), we would expect a presupposition that we can gloss as (4-2). The presupposition contains an existentially-quantified VARIABLE, which is indicated in the example by *something*:

- (4-1) a It was the bell that rang.
b What rang was the bell.
c The bell was what rang.

(4-2) Something rang.

(4-3) $\exists x \text{ rang}(x)$

We adopt a view of clefts in which a proposition such as (4-3) is both ENTAILED and PRESUPPOSED by positive clefts, and merely presupposed by negative ones.

However, the content of the cleft is not exhausted by the statement of its presupposition. We take the view that clefts also contain an ASSERTION, to the effect that the element or elements named by the cleft head serve to INSTANTIATE the variable contained in the

presupposition. We can therefore represent the remainder of the content of the clefts in (4-1) as follows:

$$(4-4) \quad \text{be}(s, x, y) \wedge \text{bell}(y)$$

where x stands for the variable in the presupposition.

Finally, a third aspect of the meaning of the cleft is that the elements specified by the cleft head—which we will argue in section 4.3 are always a (possibly singleton) set—are **UNIQUE** or **MAXIMAL** with respect to the variable in the presupposition. That is, the members of the set specified must be the only elements in the discourse context that can provide a true instantiation of this variable.

In this chapter, we propose a treatment of cleft semantics in general, and presupposition in particular, based on the view that presuppositions can be treated as **ANAPHORS** within the framework of Kamp's [1981] Discourse Representation Theory (**DRT**). This proposal, due to van der Sandt [1988, 1989], is explored in some depth, and we develop a general account of cleft semantics in **DRT** with it as a basis. In the course of the chapter, we examine a range of data in detail, and expose aspects of the cleft's meaning, several of them related to presupposition, that need to be accounted for within the final framework. In the course of the discussion, we also address several attempts to account for cleft meaning within a variety of semantic and pragmatic frameworks, evolving in the course of the discussion a view of what a representation for clefts should be like. The treatment we suggest at the end of the chapter takes these recommendations into account in providing a unified formal account of clefts in context that combines the advantages of semantic and pragmatic approaches to the meaning of sentences in discourse.

The structure of the chapter is as follows. In section 4.1, we provide the first part of the formal underpinning for the **DRT** model, providing an account of the **REFERENTIAL** aspects of cleft constructions. In this section, we show that there are important problems with one prominent account of the referential behaviour of *wh*-clefts, that of Higgins [1979]. We show that, while there is a great deal of surface diversity within and between the three types of cleft in terms of referential behaviour, a single underlying characterisation, based on an observation of Williams' [1983] with respect to copular sentences, will suffice to cater formally for the range appearing in the clefts data. We

accordingly adopt this in section 4.4 in the representation for clefts in DRT.

We move on in section 4.2 to address what is the central concern of the chapter, reflecting its crucial role in the choice and use of clefts in discourse: the nature and treatment of cleft presupposition. Having established that clefts do convey presuppositions in a fairly conventional sense, we go on to argue that the function of cleft presuppositions in discourse is to establish the content of the presupposition as information that requires an ANTECEDENT in the discourse context. If no such antecedent is available, the interpretation of the cleft requires that one be constructed. The construction of missing antecedent material is equated with the strategy of the ACCOMMODATION of presuppositions (cf. Lewis [1979]). We look through the literature in search of an account of presupposition that is able to support this view, as well as meeting other requirements for the treatment of clefts that emerge during the course of the discussion. The account we arrive at is that of van der Sandt [1988,1989], who proposes a treatment of presupposition within the framework of Kamp's [1981] Discourse Representation Theory (DRT). Van der Sandt proposes that presuppositions can usefully be viewed as ANAPHORS, adapting the existing treatment of anaphora in DRT to cater for them. At the end of the section, we provide a general characterisation of the model van der Sandt proposes.

A final aspect of the meaning of cleft constructions remains to be dealt with, however, before we explore these proposals in detail: the description and treatment of the so-called UNIQUENESS reading of clefts. In section 4.3, therefore, we go on to examine the evidence surrounding the claim that the head of the cleft specifies a discourse object or set of objects that is unique with respect to the rest of the discourse context. We suggest an analogy between this 'uniqueness' and the notion of TOTALITY described by Hawkins [1978] in relation to definite referring expressions.

Finally, in section 4.4, we examine van der Sandt's proposal within the framework of DRT, developing a treatment of cleft presupposition that draws on the treatment of pronominal anaphors. We incorporate our conclusions concerning referentiality and uniqueness into the treatment, and show how an elegant formulation of both can be arrived at. We also examine how DRT offers a framework in which some long-standing problems in presupposition can be remedied, showing in particular the relevance of the solutions it offers for the treatment of clefts.

Before we begin our development of a framework for clefts, we first have to establish some necessary terms and definitions. The largest semantic unit that we will refer to is the PROPOSITION, and each sentence that appears in a discourse realises one or more of these. Following Bach [1986], each proposition, or rather the verbal predicate it contains, is taken to describe a single EVENTUALITY, of which there are three types: STATE, PROCESS and EVENT.¹ Examples of each of these respectively (taken from Bach [1986]) are as follows:

States: *sit, stand, be drunk, be in New York, own x , love x , resemble x*

Processes: *walk, push a cart, be mean (agentive)*

Events: *build x , recognize, notice, flash once, reach the top*

In cases where more than one verbal predicate appears in a proposition, each will be considered to introduce a separate eventuality INDEX. Following the traditional Davidsonian approach to verbal predicates, the eventuality index is included as an extra argument of the predicate.

On this approach, the representation of (4-5) would be (4-6), where the index e_1 is the event index associated with the verbal predicate *eats*:

(4-5) Someone eats.

(4-6) $\exists e_1 \exists x_2 \text{eats}(e_1, x_2)$

4.1 Cleft Constructions and Referentiality

For the purposes of the DRT model of presupposition, we need to work out a coherent view of the referential behaviour of cleft constructions. In this section, we first of all draw attention to an influential, but we believe erroneous, description of referring and *wh*-clefts, due to Higgins [1979]. Higgins claims that *wh*-clefts have no referential function, in that neither the head element nor the *wh*-clause can be a referring expression. We present a range of data to show this claim cannot be supported for *wh*-cleft heads,

¹We do not use the full detail of Bach's taxonomy here, these three general categories being sufficiently fine for our immediate purposes.

and that his view confuses the issue of characterising the relationship between the cleft head and the cleft complement with independent considerations regarding the referential behaviour of the expressions that the cleft contains. We go on to show that, with respect to the expressions that appear in various positions throughout the three syntactic types of cleft, few general predictions regarding referential behaviour can be made. We argue that, on the surface, no more uniformity can be expected in cleft sentences than in other sentence types, since the presentation of a proposition in cleft form does nothing to determine the referential behaviour of the expressions within it. A generalisation can be made, however, in terms of a single underlying representation for clefts as copular sentences, based on Williams' [1983] work on the copula. The representation we propose is adopted in section 4.4 in the DRT model of clefts in context.

4.1.1 Higgins' View of *Wh*-clefts

Higgins [1979] claims that *wh*-clefts are instances of what he terms SPECIFICATIONAL copular sentences. He groups together under this heading examples such as the following:

(4-7) His height is two metres.

(4-8) His claim was that Stapleford is a great place to live.

(4-9) What I don't like about John is his tie.

Higgins argues that specificational sentences have two arguments. One, which he terms the SUPERSCRPTIONAL argument, acts as an abstract specification, such as *his height*, *his claim* and *what I don't like about John*. The other argument, termed the SPECIFICATIONAL argument (after which the sentence type as a whole is named) acts as a further specification of the superscriptional argument, an intensive complement which in effect provides a value for the other expression. *Two metres*, *that Stapleford is a great place to live* and *his tie* are all examples of such specificational expressions. Higgins explains the operation of specificational sentences in terms of two analogies: in one, the superscriptional expression acts as a variable, as we described above, the specificational expression providing a value; in the second, the superscriptional expression acts as the header of a list, with the complement supplying the list itself. With respect to the second analogy,

Higgins notes:

The heading of a list does not refer to any item at all, nor does the set of items in the list itself say anything about the heading of the list, or indeed about anything ...just as a list is neither 'about' the heading of the list or 'about' the items on the list, so, I would maintain, a Specificational sentence is neither about the Subject nor the Predicate, and therefore neither Subject nor Predicate is referential.

[Higgins 1979:213]

As we noted in the introduction to this chapter, the value/variable analogy does seem to be a good characterisation of what is taking place in cleft sentences in general. However, while such a characterisation might satisfactorily serve as a basis for a representation of the sentences concerned, there is no evidence to suggest that this value/variable behaviour has any connection with referentiality in the way Higgins suggests. For example, Higgins' view would predict that the expression *his tie* in the cleft in (4-10) is non-referential. This view requires that the expression has a different referential status from the same expression as it occurs in the declarative (4-11). The difference, however, is not at all clear:

(4-10) What I don't like about John is his tie.

(4-11) His tie is about a foot wide.

In fact, we can provide examples that show *wh*-cleft heads in particular performing a fairly normal range of referring and non-referring functions, similar to those we might expect to appear in declaratives. As diagnostics for referentiality, we can draw on tests suggested by Doron [1988] in her work on predicative copular sentences. One diagnostic provided by Doron [1988:3] is simply that if an expression has successfully referred to an element, a subsequent anaphoric reference to that same element should be possible. (4-12) shows that subsequent anaphora is possible in the case of the referent of the *wh*-cleft head in (4-8):

(4-12) What I don't like about John is his tie_i. It_i looks like a floral kipper.

A second diagnostic described by Doron states that referring expressions permit modification by non-restrictive relative clauses, while non-referring expressions do not. Recall

from chapter 2 that one of the ways in which *wh*-cleft sentences such as (4-13) can be differentiated from *wh*-relatives such as (4-14) is on the basis that the *wh*-clause of the former is non-referring, while that of the latter is not. Doron's test supports this view, demonstrating that Higgins' 'non-referential' predication is correct at least for *wh*-cleft *wh*-clauses, when taken as a constituent. The following examples show that the *wh*-relative's subject clause accepts non-restrictive relative modification, while that of the *wh*-cleft does not:

- (4-13) a *What Karen ordered, which was burnt black, was a Welsh rarebit.
 b What Karen ordered, which was burnt black, gave her an upset stomach.

However, the same test demonstrates that Higgins' specificational group do not behave consistently with respect to the referential behaviour of the 'superscriptional', subject clause. While (4-13) demonstrated that the *wh*-cleft clause was non-referential, (4-14) and (4-15) show that the subject clauses of the two declarative specificational sentences do refer:

- (4-14) His height, which is what we're concerned with, is two metres.

- (4-15) His claim, which made all the papers, was that Stapleford is a great place to live.

According to the independent diagnostics suggested by Doron, therefore, we can show that referring expressions do occur among Higgins' class of specificational sentences, which he claims to be non-referential, both in head position in *wh*-clefts, and in subject position in the declarative examples.

We would claim, contrary to Higgins' conclusions, that the use of a cleft construction of any kind does not determine the referential or non-referential status of either the arguments embedded in the *wh*-clause or complement or the arguments appearing in cleft head position. We can demonstrate this as follows, using comparisons between declarative and cleft examples. In each case, the expression embedded in the cleft *wh*-clause or complement, and the expression that appears as cleft head, appear to be playing exactly the same referential role as the corresponding expressions in the declarative examples. In fact, the range of functions that the expressions concerned can perform seem to be identical between the clefts and the declaratives. For example,

in both (4-16) and (4-17), *John* is referential, while *the baker* and *nice* respectively are predicative:

- (4-16) a John is the baker.
b It is John who is the baker.

- (4-17) a John is nice.
b What John is is nice.

The complement of the declarative, and the head of the cleft, in (4-18) are both generic expressions:

- (4-18) a We are interested in the brine shrimp.
b What we are interested in is the brine shrimp.

A similar pattern emerges in (4-19), where the declarative complement and cleft head are definitions:

- (4-19) a This is bauxite.
b Bauxite is what this is.

As a final example, clefts can carry as head both *de dicto* and *de re* elements, which can also appear as the complement of an ordinary declarative:

- (4-20) a John wants to marry a Norwegian. She has/should have long hair.
b It is a Norwegian that John wants to marry. She has/should have long hair.

The effect of the above data is to show that the range of expressions that can appear both as cleft head and embedded elsewhere in cleft constructions is apparently the same as the range that can appear in ordinary declaratives. We would claim, therefore, that the realisation of a proposition in cleft form has no effect on the referential status of the arguments in that proposition, contrary to Higgins' claim.

4.1.2 The Referential Relations of Clefts

In spite of the variety that appears in terms of referring expressions on the surface of clefts, we can postulate a single underlying representation of the referential relations in clefts. In order to do this, we abstract away from the behaviour of the surface

expressions occurring in the cleft as such, concentrating instead in what we can term the MAIN IDENTIFICATION of the sentence—that is, the identification of value with variable that we argued in the introduction of this chapter to be part of the function of the cleft. We can do this as follows.

For a cleft such as (4-21a), which yields the presupposition (4-21b), we can paraphrase the propositional content—the presupposition combined with the assertion—as (4-21c):

- (4-21) a It was John who left.
b Someone left.
c Someone_i left and it_i was John.

We can see the assertion that is part of the content of the cleft, glossed above as the second conjunct of (4-21c), as the part that identifies the variable in the presupposition with an underlying anaphoric expression of the same type. The element denoted by the cleft head, i.e., *John*, acts as the value of both that anaphoric expression and the initial variable, the copula *be* acting as the link between it and the variable expressions. The relationship between this state of affairs and referentiality is as follows. We can assume with Williams [1983] that copulas take two arguments, one referential and one predicative. It can be assumed in the majority of cases that the argument to the *left* of the copula is the referential argument, while the argument to the *right* is the predicative argument. Applying this generality to the representation in (4-21c), we can see that the anaphoric variable expression in the second conjunct is in the leftmost, referential position with respect to the copula, while the element denoted by the cleft head, *John*, is in rightmost, predicative position. We will adopt this as the appropriate general representation for clefts of all kinds, regardless of the referential status of their surface expressions. This treatment is important for the DRT model we present in section 4.4, since on the view of presupposition we present there, the existentially-quantified variable in the presupposition requires an antecedent in the context. In DRT, this antecedent has to be a discourse marker, and discourse markers are always referential in the DRT framework. Aside from notational convenience, however, the representation in (4-21c) allows us to capture a generalisation in the underlying function of the cleft which, as we have seen in the data above, has a variety of surface realisations.

Having presented our proposed treatment of cleft referentiality, we can move on to our

central concern: the nature and treatment of presupposition in clefts.

4.2 Cleft Constructions and Presupposition

Cleft constructions have for some time been considered to belong to the class of construction that induces, conveys, or requires PRESUPPOSITIONS (see, for example, Keenan [1971], Chomsky [1971], Gazdar [1979], *inter alia*).

The definition of presupposition being adopted here is broadly a Strawsonian one, in terms of the survival of entailments of the carrier sentence in negative, modal, and other contexts. There are four fairly well-known tests for presupposition: **embedding under negation**, **embedding under modality**, the so-called **antecedent test**, and the test for **constancy under illocutionary force**. According to the most traditional of the four tests, the negation test, the presuppositions of a sentence are just those entailments that are preserved under negation. Negating a sentence and comparing it with its positive counterpart should therefore reveal presuppositions as those propositions that are true in both cases. The next two tests, the modality and antecedent tests, are attributable to Karttunen [1971]; the modality test is similar to the negation test, except it uses a possibility operator rather than negation, and the antecedent test involves making the sentence under analysis into the antecedent of a conditional. In both cases, the entailments are lost, while what remains are the presuppositions. The test involving speech acts, originated by van der Sandt [1988], uses presuppositional constructions to perform a variety of speech acts; if propositional meanings are preserved across these contexts, it is assumed that these are the presuppositions of the carrier sentence. For the purposes of this discussion, success in the four tests, explained in detail below, will be taken to be the criterion for presupposition as intended here. Where other definitions of presupposition are discussed, the distinction between them and our notion will be made clear.

The Negation Test

As we noted above, the negation test shows that the proposition conveyed by the cleft complement or *wh*-clause appears to be preserved under negation. In examples (4-22)–

(4-24), the positive and negative sentences (a) and (b) both preserve the truth of (c):

- (4-22) a It is a hot-dog that she wants.
b It isn't a hot-dog that she wants.
c She wants something.

- (4-23) a What she wants is a hot-dog.
b What she wants isn't a hot-dog.
c She wants something.

- (4-24) a A hot-dog is what she wants.
b A hot-dog isn't what she wants.
c She wants something.

On the negation test, therefore, positive and negative clefts both appear to presuppose the existentially quantified version of the complement/*wh*-clause proposition. The reading of negation has to be *narrow scope* or *internal* in order for the presupposition to be preserved. Contexts which condition the broad-scope or external negation reading cause presuppositions of negative sentences to be lost. In (4-25), the embedding of the cleft causes the negative in the cleft construction to be read as broad scope within its clause, providing a reading equivalent to (4-26):

- (4-25) A hot-dog isn't what she wants, because she doesn't want anything.

- (4-26) It is not the case that she wants a hot-dog.

Since they do not rely on negation, the other three tests are not open to this ambiguity, and are therefore more reliable as tests for presupposition.

The Modality Test

The modality test involves the creation of a context in which a modal operator such as *possibly* has scope over the potentially presupposing sentence. Again, if any entailment is preserved in this context, this entailment is said to be presupposed. The complements of all three syntactic types of cleft sentence pass this test. In the following examples, the (a) sentences imply the truth of the (b) sentences:

- (4-27) a It is possible that it is a hot-dog that she wants.
b She wants something.
- (4-28) a It is possible that what she wants is a hot-dog.
b She wants something.
- (4-29) a It is possible that a hot-dog is what she wants.
b She wants something.

Antecedents of Conditionals

In this third test, as in the one before, propositions are considered to be presupposed if they are preserved in a particular context. In this case, the appropriate context is constructed by making the carrier sentence the antecedent of a conditional construction. An example of this test at work appears in van der Sandt [1988:39], using sentences with *manage to* and *be glad that*. *Manage* is an example of an implicative verb, which entails its complement. *Be glad that*, on the other hand, presupposes its complement. On this test, therefore, we would expect the truth of the complement of *manage* to be cancelled or suspended in the conditional, while the proposition carried by the complement of *be glad that* is preserved. Van der Sandt's examples show this to be the case: (4-30) does not require that (4-32) be true, while (4-31) does:

- (4-30) If Charles managed to leave the country, he will never come back.
- (4-31) If Charles was glad that he left the country, he will never come back.
- (4-32) Charles left the country.

When the antecedent test is applied to clefts, a similar pattern emerges. It appears that the cleft construction (the (a) sentence in each of the examples below) requires the complement proposition (c) to be true; the declarative (b) does not:

- (4-33) a If it is a hot-dog that she wants, Bill will be cross.
b If she wants a hot-dog, Bill will be cross.
c She wants something.

- (4-34) a If what she wants is a hot-dog, Bill will be cross.
 b If she wants a hot-dog, Bill will be cross.
 c She wants something.
- (4-35) a If a hot-dog is what she wants, Bill will be cross.
 b If she wants a hot-dog, Bill will be cross.
 c She wants something.

Constancy under Illocutionary Force

The final test for presuppositions, suggested in van der Sandt [1988:193], is that they are constant under illocutionary force. Briefly, if a proposition is conveyed by a particular sentence regardless of whether that sentence is used as a question, an assertion, or any other kind of speech act, then that proposition is an elementary presupposition of the carrier sentence. It was clear from the discussion above that clefts serving as assertions (such as (4-22a), for example) pass this test. The following examples demonstrate that clefts appearing in the form of questions (the (a) examples) preserve the truth of the (b) sentences below:

- (4-36) a Is it a hot-dog that she wants?
 b She wants something.
- (4-37) a Is what she wants a hot-dog?
 b She wants something.
- (4-38) a Is a hot-dog what she wants?
 b She wants something.

A simple declarative sentence, however, shows that entailments are not constant under illocutionary force. (4-39a) entails both *she wants a hot-dog* and (4-39c), while (4-39b) does not:

- (4-39) a She wants a hot-dog.
 b Does she want a hot-dog?
 c She wants something.

Summary

On the basis of the four tests set out above, we can adopt the traditional view that cleft constructions are presupposition-carrying in some fairly conventional sense. We will assume that the content of the presupposition can be determined by taking the content of the cleft complement or *wh*-clause and supplementing it with a suitable existentially-quantified variable. (In the examples above, we represented this informally by *someone* and *something*.) Throughout the discussion to follow, presuppositions as defined according to the four tests above will be termed the ELEMENTARY PRESUPPOSITIONS of the cleft, and the utterances which convey such presuppositions in context will be regarded as PRESUPPOSING. It is not always the case that these presuppositions remain when clefts appear in context, however: see, for example, (4-4) above. Contexts in which cleft presuppositions are lost are demonstrated at various points in the discussion to follow.

4.2.1 Characterising Presupposed Information

Having established that the cleft is a presupposition-conveying construction in a specific sense, we can turn to the examination of the kind of information that is presupposed by clefts. In particular, what is the relationship between presupposed information and the discourse context?

As we shall see in subsequent sections, presupposition can be defined according to many criteria different from our own, including what speakers believe and know, and whether this is currently shared and/or salient knowledge in the discourse (cf. Chomsky [1971], Prince [1986]); whether the context entails presupposed information (as is suggested in Karttunen [1974]); or whether it is part of the *common ground* of the discourse [Karttunen and Peters 1979]. In this thesis, we will be arguing for a view of cleft presupposition that is syntax-based, although pragmatically interpreted. The notion of presupposition argued for here does not require that at the time of utterance the presupposed proposition itself stand in any particular logical relation to the current context, nor in any single relation to hearer belief, nor that it be shared information, nor even information entailed by such information.

But why is it necessary to make these departures from the intuitive notion of ‘pre-

supposition', that of information that is presumed to be the shared background of the discourse? We take this step because it is possible to presuppose information that is not in any sense shared, known, or entailed by the context preceding the presupposing element or construction. Prince [1978] was the first to point this out in relation to *it*-clefts in the case of her INFORMATIVE-PRESUPPOSITION *it*-cleft, discussed in chapter 2. An example of this kind of cleft is as follows:

(4-40) It was just about 50 years ago today that Henry Ford gave us the weekend.

Her observation was simply that

...with [informative-presupposition *it*-clefts], not only is the hearer not expected to be thinking about the information in the *that*-clause, but s/he is not expected even to KNOW it. In fact, the whole point of these sentences is to INFORM the hearer of that very information.

[Prince 1978:898]

In fact, the phenomenon is not confined to *it*-clefts: in *wh*-clefts, as will be shown in chapter 5, a more limited range of 'informativeness' is permitted, allowing examples such as the following:

(4-41) I'm going to concentrate on dissemination and evaluation. What I haven't been able to do yet is get all the figures in.

In (4-41), the presupposition *I haven't been able to do something* is arguably not shared knowledge. Similar observations can be made for the reverse *wh*-cleft. Examples of the kinds of information that regularly appear in the presuppositions of all three kinds of cleft appear in chapter 5.

The presupposition of information that is not currently shared, salient, or standing in an entailment relation to the information currently in the discourse model is actually an instance of something that is quite general to presupposing constructions of all kinds, as has been noted by Karttunen [1974], Halliday [1985] and van der Sandt [1988, 1989], among others. On an analogy suggested by van der Sandt [1988], presuppositions are like anaphors, in that they require some ANTECEDENT in previous discourse. We would claim that a large part of the function of cleft sentences is to mark that such an antecedent is required for the presupposed information. In the frequent cases in which

the presupposed information is in fact new to the discourse, one of the effects of the cleft is to mark information as intended to be ACCOMMODATED by the hearer. That is, an antecedent for the presupposed information is intended to be incorporated into the discourse context, if it is not already available. Although the term used here is due to Lewis [1979], the phenomenon was first observed to my knowledge by Karttunen [1974] and Stalnaker [1974]. Karttunen explains it as follows:

People do make leaps and shortcuts by using sentences whose presuppositions are not satisfied in the conversational context ... I think we can maintain that a sentence is always taken to be an increment to a context that satisfies its presupposition. If the current conversational context does not suffice, the listener is entitled and expected to extend it as required. He must determine for himself what context he is supposed to be in on the basis of what was said and, if he is willing to go along with it, make the same tacit extension that his interlocutor appears to have made.

[Karttunen 1974:191]

Stalnaker makes a similar point:

A speaker may act as if certain propositions are part of the common background when he knows that they are not. He may want to communicate a proposition indirectly, and do this by presupposing it in such a way that the auditor will be able to infer that it is presupposed. In such a case, a speaker tells his auditor something in part by pretending that his auditor already knows it. The pretence need not be an attempt at deception. It might be tacitly recognised by everyone concerned that this is what is going on, and recognised that everyone else recognises it. ... I shall say that one actually does make the presuppositions that one seems to make even when one is only pretending to have the beliefs that one normally has when one makes presuppositions.

[1974:202-203]

Lewis, in his introduction to the notion of accommodation, explains as follows:

Presupposition evolves according to a rule of accommodation specifying that any presuppositions that are required by what is said straightway come into existence, provided that nobody objects.

[Lewis 1979]

The phenomenon has come to be known by several other names: van der Sandt's CONTEXTUALISATION [1982]; Seuren's BACKWARD SUPPLETION [1985]; and Reichgelt's RETROSPECTIVE UPDATING OF THE DATABASE [1986].

We would claim, then, that clefts mark the fact that antecedents have to be found or created for particular information they convey. Information marked in this way is set apart from the information that is ASSERTED by the cleft sentence, namely the information in the cleft head. On this view, the syntactic signalling of presupposition is seen as a signpost to hearers to treat information in a particular way, rather than as a signal of assumptions about the *current* state of the discourse.

Of course, not all information can be marked as intended to be contextualised. As Stalnaker points out:

The presumed background information—the set of presuppositions which in part define a linguistic context—naturally imposes constraints on what can reasonably or appropriately be said in that context.

[Stalnaker 1974]

Some of these constraints appear to be genuinely semantic, and are accordingly dealt with later in this chapter. Others, especially those that appear to be specific to the three kinds of cleft, are discussed in detail in chapter 5.

In this section, we have claimed that the function of presuppositions in general, and of cleft presuppositions in particular, is to communicate a requirement for ANTECEDENCE in a model of discourse. We would argue that such signalling is a direct consequence of the syntax of the cleft construction. In addition, because it is common to presuppose not only information that is New, but information that is currently in the context, and information that is a mixture of the two, it is plausible to postulate a notion of presupposition that is independent of any consideration of whether the information is currently present in the discourse context or not. It is therefore clear that a theory of presupposition that is going to account for this behaviour will need to distinguish clearly between information structure—what speakers share, know, believe, etc.—and the syntactic/semantic fact that a presupposition is taking place. In the following sections, we will see that few of the accounts of clefts proposed so far fulfil this requirement.

4.2.2 Accounts of Cleft Presupposition

In this section, we will discuss the major treatments of cleft presupposition that have been suggested to date. These fall into three major groups, as follows:

- accounts based on the position of sentence accent;
- accounts based on Gricean or pseudo-Gricean principles, which reduce presupposition to various kinds of implicature; and
- accounts based on the semantic interpretation of sentences augmented by a pragmatic component.

We will argue below that accounts of the third kind are the most appropriate for our purposes, and one of these in particular, that of van der Sandt [1988], suits our needs particularly well.

Accounts based on Sentence Accent

On a sentence-accent based approach, cleft presupposition is defined either wholly or partly according to the position of prosodic NUCLEI. The major proponents of this approach as applied to cleft constructions are Chomsky [1971] and Prince [1986].

Chomsky, following Halliday [1967], divides sentences into FOCUS and PRESUPPOSITION as follows:

The focus is the phrase containing the intonation center, and the presupposition is determined by the replacement of the focus with a variable.

[Chomsky 1971]

Presuppositions on this basis will be all the non-focal material in the sentence, which need not amount to a recognisable syntactic constituent. Chomsky notes that ‘there may be no actual sentence expressing just this proposition, for grammatical reasons’. Chomsky’s definition of presupposition therefore does not take the syntactic structure of

the cleft into account. Unless it can be demonstrated that this affects accentual pattern, presumably, on Chomsky's account, the cleft will be treated equivalently to any other syntactic structure.

As we will summarise in chapter 5, clefts have several distinctive accentual patterns. For the *it*-cleft in (4-42), for example, Chomsky's analysis would be (4-43):

(4-42) It was JOHN who ate the beans.

(4-43) [Focus: John] [Presupposition: x ate the beans.]

For clefts such as (4-44), however, Chomsky's definition of presupposition might produce an analysis such as (4-45):

(4-44) It was John who ate the BEANS.

(4-45) [Focus: the beans] [Presupposition: John ate x]

We would argue that Chomsky's approach takes an over-simple view of the prosodic features of sentences. First of all, as we show in chapter 5, if focus and presupposition are to be defined on sentences, many sentences contain more than one nucleus. On this basis, it is not clear how the nucleus that represents 'intonation center' is to be decided upon. Secondly, however, it seems that taking account of one principal 'center' will not capture the phenomena as Chomsky would require. Not only additional nuclei, but the appearance of subsidiary accents, are a case in point, since both appear to mark information that is arguably different in discourse status from the 'presupposed' material that Chomsky intends to isolate. As a brief example of this, we can give pairs of sentences both of which contain a presupposition on Chomsky's definition, but which in fact seem to differ in the contexts in which they are appropriate. We would claim that they differ because in the examples where a subsidiary accent appears (the (b) examples below), New information (which Chomsky might also call Focus) is present in the ostensible 'presupposition'. In (4-46) and (4-47), subsidiary accents are shown in *italics*, and nuclei in small capitals:

(4-46) A: Who did this?

a B: It was JOHN that did it.

b B: # It was JOHN that *did* it.

- (4-47) A: I'm really quite cross with you about breaking my vase.
a B: #It was JOHN that did it.
b B: It was JOHN that *did* it.

In addition, Chomsky's account assumes that focus can be reliably associated with surface lexical material. More recent approaches to capturing the notion of focus, such as that of Gussenhoven [1983], suggest that it is desirable to base focus on *semantic* material. This is because there otherwise seems to be no principled way to account for cases in which the focal item is apparently a single semantic feature, rather than a whole lexical item. For example, in (4-48), the negation, rather than the lexical material involved, might arguably be termed the focus:

- (4-48) A: They should clean up after themselves occasionally.
B: But they DON'T!

Even on an intuitive basis, the cases where subsidiary accents occur seem to be closer to Chomsky's 'focus' than his 'presupposition', yet his account does not capture them.

Chomsky's account, as is common to accounts based on sentence accent, takes the view that presupposition is a phenomenon of the same order as what is sometimes termed INFORMATION STRUCTURE—the oldness, newness, or retrievability of information in the discourse from the hearer's point of view, rather than a syntactic/semantic phenomenon. On the account being developed here, these levels are separated out, sentence accent being associated with the latter, and syntax with the former. Sentence accent accounts of presupposition do not allow for the observation that presuppositions appear to be systematically related to syntactic structure.

While at this point to adopt a syntax-based account of presupposition may seem to be a matter of simple preference, we will see shortly that the accentual treatment of presupposition is unable to deal with any of the problematic cases for presuppositional theory, in particular those in which presuppositions are lost while sentence accent remains unaltered. In fact, this difficulty is common to both Chomsky's account and another prominent account of presuppositional phenomena in pragmatically-marked syntactic structures, that of Prince [1986]. Before going on to examine the problems in the approach, therefore, we will first of all outline Prince's claims.

Prince's [1986] view of cleft presupposition, although still based on sentence accent,

does allow the syntax of the cleft to play a role in determining presupposition. She argues that *it*-clefts and *wh*-clefts fall within a class of sentence types that can mark variable-containing or OPEN PROPOSITIONS (OPs) as information presupposed by the speaker; i.e., information taken to be salient shared knowledge in the discourse. OPs are of course equivalent to the kind of standard existentially quantified proposition we have already isolated as the elementary presupposition of clefts.

On Prince's account of *it*-clefts and *wh*-clefts, presupposition is suspended or cancelled just in case the prosodic nucleus of the sentence appears in the part of the sentence that contains the putative presupposition, namely the cleft complement or *wh*-clause. A presuppositional reading occurs when the gap-binding constituent (in the cleft's case, the cleft head) carries the nucleus,² but presupposition is cancelled in cases where new material—and hence the nucleus—appears in the potentially presupposing part of the construction.

It is interesting to note, too, that on Prince's account, *wh*-clefts are always OP-marking, while *it*-clefts need not be. This view of *wh*-clefts requires the assumption that nuclear accents do not occur in *wh*-cleft *wh*-clauses. As our study of the Survey of English Usage data presented in chapter 5 shows, nuclei frequently appear in this position (in the sample of 50 *wh*-clefts used in the study, 41 had nuclei in both the head and the complement). An example of such a *wh*-cleft is as follows:

(4-50) What they KNOW about is experimental RESEARCH.

Approaches such as Prince's and Chomsky's seem to capture some intuitive notion of what speakers are assuming to be shared knowledge at the time of the utterance. However, in their neglect of syntactic structure, they cannot explain the additional and complementary marking system that is based on syntax. For example, Chomsky's

²Prince does, however, comment that OP-marking still takes place even when the gap-binding constituent isn't the one bearing the nucleus in cases where the gap is of category VP or dominated by VP. This would predict that reverse *wh*-clefts with VP heads, such as

(4-49) Making jigsaws was what he was doing

would be categorised as OP-marking, regardless of accentual pattern. This does not seem to be quite correct, since, as we shall see in chapter 5, the information structure of these clefts is as free as that of the *it*-cleft—that is, they are able to carry new information in the complement, which cannot be described as 'shared' and 'salient'.

sentence accent account would analyse the presupposition of both (4-51) and (4-52) to be (4-53):

(4-51) At four I'm going to SLEEP.

(4-52) It's at four I'm going to SLEEP.

(4-53) I'm going to do x at four.

However, the analysis cannot account for the differences in meaning between of the two sentences in a context such as the one that appears in (4-54):

(4-54) A: Have you got anything planned for this afternoon?

a B: At four I'm going to SLEEP.

b B: It's at four I'm going to SLEEP.

The cleft example, (4-54b), suggests that speaker B is *reminding* speaker A of her plans for the afternoon, giving what we might informally term an *as you know* reading. If this interpretation is not available (for instance, if A is a person from the electricity board that B has never seen before), (4-54b) is unacceptable as a response. The declarative simply informs A of B's plans, and does not suggest a previous acquaintance with the knowledge.³

As we saw, Chomsky's sentence accent account gives an identical analysis for both, while Prince's would simply cancel the presupposition in the cleft example. Neither account captures the extra meaning attached to the syntax of the cleft, which we have claimed is due to the requirement of a contextual antecedent for the syntactically-defined presupposition for the cleft. The notion of presupposing new information is on both accounts highly problematic, since such information would have to be unaccented in order to pass as presupposed. This would mean that any algorithm developed to differentiate focal from presupposed information would have to distinguish new information that *falls within the scope* of an accent yet is not itself accented⁴ from new information that is deliberately unaccented in order to pass as presupposed.

³As we will see in chapter 5 in our discussion of presupposition in context, prior acquaintance with the information carried by the cleft presupposition is by no means a requirement for the cleft to be acceptable in discourse. On this occasion, the '*as you know*' reading can be accounted for by the high degree of intersection that exists between the information that is signalled to be genuinely available to the hearer by an accent pattern such as that in (4-54b), and the information marked syntactically as requiring an antecedent.

⁴Not every syllable of 'New' information can be accented, as noted by Halliday [1967]. It is a well-

Most importantly, however, theories such as these cannot accurately predict circumstances in which presuppositions will fail to survive. It is well known that presuppositions can be cancelled or suspended in certain situations, depending on context. This problem, known as the PROJECTION PROBLEM for presupposition, resides in the fact that presuppositions of simple sentences be suspended or cancelled under complex embeddings, for example in conditional contexts. This is true even of sentences that presuppose on Chomsky's and Prince's definitions. (4-55a), for example, fulfils their criteria, and as a result presupposes (4-55b):

- (4-55) a It was JOHN who ate the beans.
 b Someone ate the beans

In (4-56), the same example appears as appears as the consequent of a conditional. The presupposition in this case can only be said to survive if the antecedent clause of the conditional is true:

- (4-56) If anyone ate the beans, it was JOHN who ate them.

On Chomsky's and Prince's accounts, however, the presupposition would remain, since the sentence has the appropriate accentual structure. This leads to an incorrect analysis of (4-56) as presupposing, and would do the same for any sentence which appears in a context in which presuppositions fail to project. The situation is the same for examples such as (4-57), where a broad-scope negation suspends the presupposition without affecting accentual pattern:

- (4-57) It wasn't John who ate the beans, because nobody did.

This demonstrates the necessity for separating the level of the actual status of information with respect to a hearer's model of the discourse, which is what accounts such as Chomsky's and Prince's capture, from the pragmatic interpretation of sentences in contexts. It is our view that the notion of shared knowledge is suitably captured at the level of INFORMATION STRUCTURE, that is, by describing the relationships that exist between the content of the cleft and the discourse context at the time of the cleft's

known difficulty both for research in pragmatics and in phonological accounts of sentence accent that a single accent appears to mark a larger stretch of the discourse than the accented syllable itself. For approaches to how to deal with the scope of accents, see in particular Selkirk [1980], Gussenhoven [1983], and Taglicht [1984].

interpretation. Considerations such as this are often relevant in the pragmatic interpretation of presupposition, but on our view, the generation of presuppositions resides at the level of syntax.

Accounts based on Implicature

Several attempts have been made to account for presuppositional phenomena in terms of Gricean implicatures, both conventional and conversational.⁵ The major account of clefts that falls within this category is that of Atlas and Levinson [1981], who argue that conversational implicature is central to the description of *it*-clefts.

Atlas and Levinson argue that an important part of the ‘meaning’ of clefts is non-truth-conditional, a view that is compatible with the position we argue for later in this chapter. However, their account differs from ours with respect to the explanation they give regarding the *source* of this non-truth-conditional meaning. Most phenomena that are currently attributed to presupposition, they argue, can in fact be explained equally as well by a combination of entailment and conversational implicature. Their argument is that, instead of clefts carrying presuppositions, positive clefts (which can also be termed *affirmative* clefts) carry entailments, while negative clefts carry implicatures in addition to entailments. For example, the positive cleft in (4-58) is claimed to entail (4-59) and (4-60), but (4-59) does not entail (4-58):

(4-58) It was John that Mary kissed

(4-59) Mary kissed John

(4-60) Mary kissed someone

The notion of implicature is relevant in the description of negative clefts such as (4-61). Atlas and Levinson claim that this example would entail (4-62) (although (4-62) does not entail (4-61)), and implicate (4-63):

(4-61) It wasn’t John that Mary kissed

⁵A full critique of accounts of this nature appears in van der Sandt [1988:50ff]; in this study, we concentrate on accounts which have addressed cleft constructions specifically.

(4-62) Mary didn't kiss John

(4-63) Mary kissed someone

The argument Atlas and Levinson present in support of the analysis of (4-63) as an implicature rather than a presupposition is as follows. In order for (4-63) to be interpreted as part of the meaning of (4-61) at all, regardless of whether it is presupposed or implicated, the negation in the example has to be interpreted as INTERNAL or NARROW SCOPE. This much is uncontroversial. Along with Allwood [1972], Gazdar [1976, 1979] and others, Atlas and Levinson claim, however, that the narrow scope reading is not the natural or 'default' interpretation of negation. Atlas and Levinson argue that, for a proposition such as (4-63) to be part of the meaning of the cleft, some principle has to operate to guide the reader or hearer to choose the non-default, internal, negation. The natural reading is the external negation, which, for (4-61), would produce a reading that can be glossed as (4-64):

(4-64) It is not the case that Mary kissed John.

To ensure that the internal negation reading is assumed where appropriate, Atlas and Levinson posit a pragmatic principle to condition the selection of internal negation on the grounds that it yields the most INFORMATIVE interpretation of the sentence. They term this principle the PRINCIPLE OF INFORMATIVENESS. Atlas and Levinson argue that this inference to the 'best interpretation', that is, from external to internal negation, is a generalized conversational implicature. The fact that the process is mediated by a (quasi-Gricean) principle, rather than performed on the basis of context or the structure of the sentence, is what leads Atlas and Levinson to maintain that the proposition in question is implicated rather than presupposed.

Several difficulties are apparent in Atlas and Levinson's account, however. Firstly, it will be recalled that their analysis states that the proposition conveyed by the complement of a positive cleft is simply an entailment, while that of the negative is implicated. The first of these two claims seems to be beyond doubt. Evidence for this is the fact that the behaviour of this proposition when cancellation is attempted is typical of entailments. (4-65) demonstrates that, like entailments, the proposition cannot be cancelled: if the sentence in (4-65) is true as a whole, then the proposition in (4-66) cannot be false:

(4-65) *It was John who ate the beans, although nobody did.

(4-66) Someone ate the beans.

However, the tests presented at the beginning of this section demonstrated that propositions like (4-66) are also presupposed. Further evidence for their presupposed status is given by van der Sandt [1988:77], who points out that the ‘entailment-only’ analysis does not explain why readers and hearers would continue to infer the truth of the proposition concerned even in non-entailing environments. Non-entailing environments include, once again, conditional contexts. In such contexts, no explicit negation occurs, but entailments of other entailing constructions (such as declaratives) fail to be preserved. In (4-67a), the declarative antecedent of the conditional, *John ate the beans*, loses its entailment; but in (4-67b), which is a cleft in a similar context, preserves the proposition *someone ate the beans*, carried by its complement:

(4-67) a If John ate the beans, Bill will be cross.

b If it was John who ate the beans, Bill will be cross.

To account for these facts, the only interpretation open is one to the effect that the proposition conveyed by the cleft complement is presupposed as well as entailed. The analysis of the positive cleft as simply entailing, therefore, obviously will not do. In fact, as we will see in section 4.3, Atlas and Levinson ultimately find it necessary to appeal to the notion of presupposition in positive clefts even though they discount it here, in order to account for a related aspect of cleft meaning.

For negative clefts, Atlas and Levinson’s claim is that the proposition conveyed by the cleft complement is implicated, rather than entailed or presupposed. It is quite clear that no entailment is involved in this case, since cancellation is possible:

(4-68) It wasn’t John who ate the beans, because nobody did.

At the beginning of this section, we showed that negative cleft complements, too, conveyed a presupposition. However, it is also possible to show—and this is the substance of Atlas and Levinson’s claim—that negative cleft complement propositions also satisfy the standard Gricean definition of implicature; that is, the implicated proposition is non-detachable and cancellable.

We can demonstrate this as follows. First of all, detachability requires that the implicature-generating sentence in question can be replaced by another expression in the same context with the same broad meaning, and still generate the implicature. (4-69) demonstrates that, in context (a), if the negative cleft (b) is replaced by another sentence (c) (a declarative, in this case), the resulting discourse will cease to convey proposition (d):

- (4-69) a I thought I'd explain about the state of the plates.
b It wasn't John who washed up.
c John didn't wash up.
d Someone washed up.

The two readings are quite distinct: (4-69b) preserves the implication that someone had washed up (presumably badly); (4-69c) loses any such implication, resulting in the interpretation that no-one had washed up at all. The example as a whole demonstrates that (4-69d) is a NON-DETACHABLE part of the meaning of (4-69b).

Cancellability requires that, when the sentence carrying the putative implicature is placed in a different context, the implicature disappears. Such cancellability appears in clefts, as can be demonstrated by a context which forces an EXTERNAL or WIDE-SCOPE reading for the negation, such as that in (4-70):

- (4-70) It wasn't John who washed up, because nobody did.

It appears, then, that the proposition conveyed by the cleft complement conforms both to the definition of implicature, and, as we showed above, to our definition of presupposition. It might then appear that the two analyses of negative clefts do not appear to differ in any crucial sense. We might, however, question the basis for Atlas and Levinson's claim that implicature, rather than presupposition, is the relevant term to apply because of the explanation they give for the source of the 'implicated' proposition, namely the Principle of Informativeness. It will be recalled that the Principle is required to license the inference from the external negation that Atlas and Levinson see as the default, to the implicature-bearing internal negation. However, as Givon observes, the assumption that the default negation in natural language is external may not bear close inspection:

In general, while linguists and philosophers find it easy to recognise the ex-

ternal sense of negation, most speakers of human languages do not. In other words, they tend to view negative constructions almost always as internal operations.

[Givon 1978:88]

If Givon's claim is true, the basis for Atlas and Levinson's argument that implicature is taking place—the inference from the 'default' external negation to the required internal negation via the Principle of Informativeness—is flawed.

In what follows, we will continue to use the term PRESUPPOSITION to describe the propositions conveyed by the complements of negative clefts. This is partly because of the objection raised above, and partly because, as van der Sandt points out, there seems little point in invoking specialised pragmatic principles to explain a notion that is more simply and generally captured at the level of semantics. For positive clefts, however, our conclusions differ from Atlas and Levinson's in more than a cosmetic sense, since we have shown that both entailment and presupposition are taking place. The evidence from non-entailing contexts shows that the propositions in question do more than merely entail.

Semantic/Pragmatic Accounts

In this last broad category of treatments of presupposition reside theories that, in general, separate out the level of semantic interpretation of sentence types from the process of interpreting the resulting intermediate structures pragmatically, in order to produce the final representation of the sentence in context. Into this category falls the final theory of presupposition to be discussed, namely that of Gazdar [1979].

Gazdar's proposed solution to the projection problem is to build up an initial representation of the component simple clauses of sentences to be processed. From this representation, all the *potential* presuppositions and implicatures for each clause are generated. The potential presuppositions (PRE-SUPPOSITIONS, in Gazdar's terminology) and potential implicatures (IM-PLICATURES) need to be filtered in some way to see which are to survive in order to represent the sentence meaning correctly. This is achieved by a mechanism that causes all the propositions in the sentence to be integrated

into a discourse context, itself a set of propositions, in a certain order, the unwanted ones being discarded under conditions of inconsistency. The order for integration is as follows:

- entailments
- clausal im-plicatures
- scalar im-plicatures
- pre-suppositions

The condition for the integration of the representation of any clause is simply that it should be consistent with the context. On Gazdar's account, both *it*-clefts and *wh*-clefts fall into the group of sentence types for which a function can be defined to yield a pre-supposition. While this function is not given explicitly, Gazdar comments that it would fit the following general pattern :

...clefts, pseudo-clefts and *wh*-questions ...share a pre-supposition that is arrived at, more or less, by extracting the *wh*-clause and substituting the appropriate existentially-quantified phrase for the *wh*-phrase within it.

[Gazdar 1979:128]

The pre-suppositional analysis of an *it*-cleft such as (4-71) would therefore be (4-72), that of the *wh*-cleft (4-73), (4-74), and by analogy (although Gazdar does not mention these explicitly) that of a reverse *wh*-cleft such as (4-75), (4-76) (the *speaker knows* gloss follows Gazdar's analysis):

(4-71) It was Duane who owned all those halter tops.

(4-72) Speaker knows that someone owned all those halter tops.

(4-73) What finished it was the black patent pumps under the bed.

(4-74) Speaker knows that something finished it.

(4-75) That was what made Pandora give all his clothes to the PDSA.

- (4-76) Speaker knows that something made Pandora give all his clothes to the PDSA.

The presuppositions of a sentence are those pre-suppositions that survive the process of integration. Assuming that no propositions already exist in the knowledge base that will cause the pre-suppositions not to survive on the grounds of inconsistency, it can be assumed that they will appear as presuppositions in the final analysis of the sentence types.

Gazdar's [1979:123] approach to the initial analysis of cleft sentences is similar to the one we argued for in the previous section, in that he supposes positive clefts to entail as well as presuppose the proposition conveyed by the complement, while negative clefts only presuppose. His analysis of the positive cleft in (4-77), and the corresponding negative cleft in (4-81), would therefore be as follows. (4-77) would be considered to entail (4-78) and (4-79), and presuppose (4-80):

- (4-77) It was the woodlice that ate the nasturtiums.

- (4-78) The woodlice ate the nasturtiums.

- (4-79) Something ate the nasturtiums.

- (4-80) Speaker knows that something ate the nasturtiums.

(4-81) would entail (4-82,) and presuppose (4-83):

- (4-81) It was not the woodlice that ate the nasturtiums.

- (4-82) The woodlice did not eat the nasturtiums.

- (4-83) Speaker knows that something ate the nasturtiums.

Although we would agree with this analysis, difficulties arise with Gazdar's model for clefts with respect to the mechanism he proposes for integrating propositions into context. Before striking this negative note, however, we should observe the positive advantages of the mechanism, which does allow Gazdar to avoid some of the anomalies that would otherwise arise in the case of clefts. For example, Keenan [1971:52] gives an example in which clefts are iterated (cited in Gazdar, [1979:149]) to demonstrate the falsity of the pre-supposition of the cleft:

- (4-84) You say that someone in this room loves Mary. Well, maybe so. But it certainly isn't Fred who loves Mary. And it certainly isn't John ... (We continue in this way until we have enumerated all the people in the room). Therefore no one in this room loves Mary.

Although each cleft in itself pre-supposes the proposition *Someone in this room loves Mary*, Gazdar's model would allow the initial sentence *You say that someone in this room loves Mary* to implicate that it is possibly not the case that someone in the room loves Mary, therefore (since implicatures are integrated before pre-suppositions) preventing the pre-suppositions of the clefts that come after from being integrated.

The ordering of integration also allows Gazdar's model to deal with a situation pointed out by Wilson ([1975:123], given in Gazdar [1979:141]) in which a *wh*-cleft yields an entailment and a pre-supposition that are contradictory. Wilson shows that (4-85) yields the inconsistent propositions (4-86) (the entailment) and (4-87), a presupposition (pre-supposition, in Gazdar's terms):

- (4-85) What your generalization captures is precisely nothing.

- (4-86) Your generalization captures nothing.

- (4-87) Your generalization captures something.

Again, Gazdar's approach would cause (4-87) to be discarded, since it is incompatible with the proposition integrated before it.

The difficulties with the integration mechanism referred to above arise precisely because the preferred analysis for positive clefts is to have the same proposition presupposed as well as entailed. While there is no doubt that this is the correct analysis, Gazdar's model loses the distinction between this and an entailments-only analysis (such as might arise in the case of a declarative), since entailments are integrated before presuppositions. The identical presupposed proposition is therefore redundant by the time it is processed. Since this is equivalent to treating the presupposition as an entailment, the model fails to capture any notion of what it *means* to presuppose with respect to a model of discourse.

In addition, it means that the integration of 'bad' presuppositions—those that are anomalous for pragmatic reasons, for example—cannot be constrained. As Prince [1978] has shown, and as we examine further in chapter 5, what is acceptable as an entailment

of a declarative is not always acceptable when presented in a presuppositional construction:

- (4-88) a Hi! My name is Ellen.
b *Hi! What my name is is Ellen.

Since presuppositions and assertions cannot be distinguished, and the sole criterion for the acceptability of a proposition is its consistency with that context, pragmatic anomaly such that occurring in (4-88b) cannot be prevented—the reason for its lack of acceptability is presumably not inconsistency.

Finally, if we are interested in the treatment of discourse in a broad sense, we should note that because the integration mechanism Gazdar proposes fails to integrate inconsistent propositions, it cannot cope with the processing of contradictory statements that we might expect to appear in dialogues.

A general criticism of models in which presuppositions are cancelled or suspended in case of inconsistency (or, on models such as Prince's, in case new information appears in the presupposed part of the utterance) is that it is not possible on these accounts to support the claim that presuppositional constructions have a pragmatic function in discourse *even in cases where presuppositions would traditionally be suspended*, such as the conditional contexts that we have discussed above. We have claimed above, and will discuss further in chapter 5, that presuppositional constructions convey processing signals to hearers and readers, requiring that antecedents be found for the presupposed material. In cases where presupposition is suspended, it is not the case that these signals disappear and cease to be processed: in other words, the rhetorical force of the cleft is not lost. On the account of cleft presupposition that we adopt in this thesis, allowance is made for these situations by the postulation of LEVELS OF EMBEDDING in the discourse model in which the antecedents are to be constructed which are used for the interpretation of presuppositional structures whose presuppositions do not ultimately survive. This will be discussed in some detail below: for the moment, it is sufficient to note that models that cancel presuppositions in these cases, or fail to differentiate between presupposed and entailed presuppositions even when presuppositions do survive, lose vital information concerning how the content of the sentences concerned are to be integrated into context. While Gazdar's analysis of the cleft is in our view

correct, therefore, the mechanism he proposes for the integration of the cleft's content into context is not sufficiently sensitive to allow pragmatic considerations to play a part in the interpretation. This sensitivity—to allow factors such as appropriateness to come into play, for example—is a vital feature of any model that is to capture the discourse meaning of presuppositional constructions in general, and clefts in particular.

4.2.3 Requirements for a Theory of Cleft Presupposition

On the basis of the discussion above, what features does a good theory of presupposition for clefts need to have? In this section, we summarise conclusions from our discussion of the three types of presuppositional theory, and augment the summary with comments on additional features that the clefts data demonstrate to be important.

How Presuppositions should be Generated

In the preceding sections, it was argued that a suitable view of presupposition should take the syntax of the construction in question into account, since it is clear that presuppositional behaviour in the cleft is due in no small part to its syntactic structure. A theory would therefore be required that allowed at least *potential* presuppositions to be generated in a manner based on syntactic structure. Accent-based accounts of presupposition, it was argued, neglect the important fact that the syntactic structure of the cleft is itself presuppositional, in addition to confusing the semantic level of description at which presupposition operates with the pragmatic issues surrounding how presuppositions cohere with the discourse contexts in which they are used.

It is also important that the means of generating potential presuppositions should be based on a well-grounded view of how negation works. It appears that for the cleft at least that the unmarked case for negation is internal, rather than the default external negation posited by approaches based on the Gricean pragmatic view.

Finally, an adequate theory needs to generate presuppositions for both positive and negative clefts, at the same time capturing the asymmetry between them revealed by cases of presupposition 'cancellation'. As we argued above, positive clefts entail and

presuppose the existentially-quantified content of the cleft complement, while negative clefts only presuppose.

What Context Should be Like

What kinds of propositions can make up the context is vitally important. We argued in the discussion of notions of presupposition based on sentence accent that the propositions that the discourse context would be likely to contain need to be considered separately from the knowledge and beliefs of the speakers. The discourse context might contain information from such diverse sources as the following:

- Propositions that have been introduced explicitly in the conversation.
- Background information shared by both speakers.
- Propositions adopted by speakers for the sake of argument.
- Propositions inferred or otherwise indirectly derived from others explicitly introduced.

Propositions in the context need not be shared knowledge, or believed to be true by either speaker (and in fact can be believed to be false by either or both).

Finally, it may be obvious that the context will need to change dynamically, to take account of propositions being integrated and, on occasion, retracted.

How Propositions Should be Integrated into Context

It was argued above that a mechanism that allows propositions to be integrated into the context under the sole constraint that they are consistent with what is already there is too simple for our purposes. Consistency is naturally important, but other, more subtle constraints are involved if *discourse* constraints on the presentation of information are to be observed in addition to logical ones. The mechanism should separate out presuppositions, assertions, entailments and implicatures as distinct cases so that pragmatic constraints on the presentation of information in each of these forms can be

implemented. Moreover, any component of the integration mechanism that evaluates or constrains integration on the basis of information structure (i.e., 'old' versus 'new' information) should make a sharp distinction between this information structure and the asserted, presupposed, entailed or implicated status of the proposition that carries it. On this basis, assessments of 'conversational contribution', if provided, can cope with new information being presented as presuppositions, for example. Most importantly for our purposes, this provision also enables the distinctive discourse effects of presupposing new information to be captured. The integration mechanism should be able to assess the acceptability of repeated information, and redundancy in the presentation of information, in a variety of situations, to ensure that utterances that violate Gricean maxims are not accepted automatically. This latter stipulation is particularly important if the mechanism is ever to be extended to deal with implicatures that *result* from such violations.

As we noted above, in models where presupposition can be cancelled or suspended (as might happen, for example, when the presupposition occurs in the consequent clause of a conditional construction), the information that a presupposition has taken place should not be lost. This reflects the fact that presupposition-signalling constructions are used deliberately and systematically, whether the presupposition itself projects or not. A model for processing presuppositions would be able to capture this most appropriately by reflecting the signalling role of presuppositions in the way in which it integrates them into the discourse context. If possible, the model for integration should be compatible with extensions to deal with dialogue.

General Requirements

It is to be hoped that a framework appropriate for clefts is generalisable to other presupposing constructions, as well as being elegant and, hopefully, simple.

In the following section, we will introduce a model, due to van der Sandt [1988, 1989], that displays most of the desirable characteristics which we argued above were necessary for an adequate account of cleft presupposition.

4.2.4 A Model for Presupposition

Van der Sandt's [1988, 1989] model for presupposition centres upon the following points:

- Presupposition is not a binary relation between sentences and propositions; it is a ternary relation between sentences, contexts and presuppositions. On this view, the presuppositions of sentences are context-sensitive, and cannot therefore be determined out of context.
- The presuppositions of a syntactic structure or lexical item act as functions that *select* from a set of possible contexts those contexts in which they are appropriate. Presuppositions *require*, as a condition for interpretability, a context in which the presupposed proposition has already been asserted or is assumed to be true. If the appropriate context is not given, the hearer can postulate it for the purposes of interpreting the sentence [van der Sandt 1988:163].
- Contexts change as a result of the utterance of new sentences.
- With most speech act theorists, we assume that the objects of speech acts are propositions, and that with each context is associated a set of propositions taken to be true in that context.
- A discourse or conversation must be seen as a sequence of speech acts and a text as a representation thereof.
- Acceptability does not affect the truth conditions of the sentence; the rules determining acceptability fall within the scope of pragmatics.
- Presuppositions are viewed as ANAPHORS, in that they require satisfaction in the context
- Drawing on insights from Discourse Representation Theory (henceforth DRT), the context permits separate LEVELS OF EMBEDDING which constrain the accessibility of antecedents depending on the depth of embedding of the 'anaphor' (i.e., the presupposition). This enables presuppositions that do not survive in a particular context to be represented, thus preserving parallels between the processing of both non-projecting and projecting presuppositions

In this section, these principles will be amplified in sufficient detail to demonstrate the usefulness of van der Sandt's model for clefts. We then go on to pursue van der Sandt's suggestions in the form of a DRT model of cleft presupposition, which we present in section 4.4.

The Context-Sensitivity of Presuppositions

Although we have established that clefts are presuppositional constructions, suspension of these presuppositions can take place in a variety of contexts, some of which we have already noted in this chapter. Suspension may occur in conditionals, under verbs of propositional attitude and verbs of reported speech, and in sentences containing temporal clauses. Van der Sandt notes, however, that suspension of presupposition not only depends on the linguistic context, but on *the assumptions made by the interlocutors*. He points out in support of this claim [1988:158] that not *all* conditional contexts cause presuppositions in the postcedent clause to be cancelled, depending on what is tacitly assumed by the speakers to be part of the context. For example, (4-89) allows both a presuppositional and a non-presuppositional reading depending systematically on the contextual information selected—that is, depending on what information is in the context at the time of utterance.

- (4-89) If someone at the conference solved the PROBLEM, it was Julius who solved it.

In the antecedent clause of (4-89), what is at issue is whether or not the problem was solved. In this case, what is in question matches the content of the presupposition of the cleft in the consequent clause, and the example as a whole acts as a conventional conditional context. In this case, therefore, as in ordinary conditional contexts, a non-presuppositional reading is available for the cleft depending on whether or not the antecedent clause is decided to be true or not. In the antecedent clause of (4-90), however, things are somewhat different. Here, what is at issue is only *where* the problem was solved. In this case, what is in question does not affect the truth value of the cleft in the consequent clause in the same way: whether or not the antecedent clause is true, the cleft presupposition will still be true. That is, the problem appears to be solved whether or not Julius solved it:

- (4-90) If someone at the CONFERENCE solved the problem, it was Julius who solved it.

These phenomena suggest that the computation of the elementary presuppositions of sentences have to take into account not only the syntactic form of the sentence in which they appear, but more sophisticated features of context such as shared knowledge. As van der Sandt points out, this evidence points towards a definition of presupposition as A THREE-PLACE RELATION between sentences, propositions, and contexts.

The Nature of the Context Set

We have already said something about the nature of the ideal discourse for our purposes. The discourse context as conceived of by van der Sandt [1988:164] has the following features:

- It contains background information that is shared by the interlocutors
- It contains information that has been introduced in the course of the conversation, provided that it is accepted and not refuted or withdrawn
- It can be extended by non-linguistic events that are open to perception by all participants
- It continually changes throughout the discourse, due to the constant addition (or retraction) of information

Propositions in the context set do not need to belong to common knowledge, or be believed by the interlocutors. Neither do the interlocutors' views about which of the propositions in context is background knowledge and which has been asserted have to coincide with one another: on this model, it is possible for different speakers to have different perspectives on the content of the context set. It will be noted that Gazdar's [1979] theory only caters for explicitly introduced information: that is, only that information finds its way into the common ground. It is therefore not possible to use such a notion of context to cater for sentences that presuppose tacitly assumed information.

Sentence Interpretation Depends on Contextualisation

In this and the following section, we embark upon topics that are particularly useful for an understanding of the rationale behind the DRT model of presupposition we present in section 4.4.

Van der Sandt describes the role of presupposition in contextualising information as follows:

The interpretation of sentences ...often requires that we construct contexts in which those sentences are acceptable and interpretable. Furthermore presuppositions are normally supposed to belong to those contexts. Given these two facts we can understand the role of elementary presuppositions in determining the presuppositions of sentences in isolation; they can be conceived of as syntactic or lexical indicators for the selection or construction of contexts. A presupposition theory can then be conceived of as a set of principles for selecting or constructing partial contexts.

[van der Sandt 1988:181]

Each presupposed proposition therefore selects a set of contexts in which it is appropriate, each of which contain the proposition itself. Each time the presupposing sentence is used, it has to be in one of these contexts.

The constraints on what is acceptable in the context will be discussed shortly: in the meantime, it is important to understand the contribution of discourse representations to van der Sandt's model. Several versions of Discourse Representation Theory exist (see, for example, Kamp [1981], Heim [1982]); it is sufficient for our purposes to communicate a common feature which makes DRT particularly suitable for an account of presupposition as van der Sandt describes, namely the hierarchical ordering of the discourse representation structures that are built up during processing of a sentence or sentences. A representation is built incrementally for simple sentences, but when logical connectives, conditionals, etc. are encountered, more complex embedded structures result. Conditionals, for example, induce two structures, where the one which represents the consequent is subordinated to that representing the antecedent. Each representation, whether subordinated or not, consists of the objects that have been introduced as the discourse is processed (usually termed DISCOURSE MARKERS) and predicates and

relations of those objects. We will show how clefts can be represented in this framework shortly. For the moment, we should note the features that make it particularly suitable for our purposes.

What is important for this account is the way the hierarchically-constructed discourse representation is able to constrain anaphoric relations between entities existing in its member structures. A discourse marker can act as an antecedent to an anaphor only if it is contained in the *same* structure as the marker of the anaphor or if it is on a superordinate level of structure. Van der Sandt [1989:21] suggests that presuppositions work similarly to anaphors, meaning that the presupposing material has to be in a well-formed anaphor-antecedent relation with appropriate material elsewhere in the structure. How this works for cases of presupposition projection is explained in van der Sandt [1989:12ff]; particularly relevant for our purposes here is to see how cases of accommodation are dealt with in the model. When no antecedent can be found for a presupposition, a discourse marker is created in the structure to act as antecedent, taking advantage of the fact that presuppositions of this nature have sufficient descriptive content to allow such antecedents to be formed. The marker will be placed at the highest level in the structure possible without violating the well-formedness conditions on structures internal to DRT. When discourse markers created in this way are able to appear at the top level of the representation, these are presuppositions of the sentence.

The important thing to note about the model is that antecedents are created for *all* presuppositions, regardless of whether or not they belong to the final set of presuppositions of the sentence. Instead of suspension or cancellation, the DRT model reflects the fact that presupposition has occurred, capturing the circumstances in which presuppositions fail to project by using constraints on the structure of the model being built. If pragmatic effects arise from both projecting and non-projecting presuppositions, as we have claimed they do, the ability of the model to capture both makes it a particularly useful basis for a pragmatically-conditioned model of for discourse processing.

Constraints on Contextual Acceptability

Van der Sandt's model, as described above, treats propositions as instructions to construct appropriate contexts for their own interpretation. In order to constrain the incorporation of sentences into contexts, as well as the formation of antecedents for presuppositions, van der Sandt [1988:199] suggests four principles, which can be paraphrased as follows:

Consistency: As is common in models where new propositions have to be integrated, propositions that are included in the context have to be consistent with it.

Informativeness: The information contributed by a sentence must not be a subset of that already in the context.

Efficiency: This is a variant of the previous principle, ensuring, for example, that the antecedents of conditionals do not already exist in the discourse context. It too helps to prevent redundancy.

Sequential Interpretation: Sentences are to be processed left to right.

The principle of consistency prevents inconsistent sets of propositions arising, since such sets entail everything. The principle of informativeness ensures that propositions cannot simply be re-iterated, and incorporated into context *ad infinitum*. This will be particularly important in the next chapter, where we will look at the constraints imposed by the context on what information can be presented in cleft form. The principle of efficiency prevents unnecessarily complex utterances that may give rise to implicatures that less complex structures would not. Finally, the principle of sequential interpretation ensures that the correct DR structures can be built up. This allows conjunctions, disjunctions, and conditionals to be processed correctly, with the appropriate subordination relationships between the representations of their parts, and between their parts and the existing structures in the context.

In section 4.4, we show how this general description relates to a treatment of cleft presupposition in particular. Before we turn to this, however, there is a final aspect of the meaning of clefts to be examined: the 'uniqueness' reading for cleft heads.

4.3 Cleft Constructions and Uniqueness

In this section, we will examine the claim that clefts implicate, presuppose or otherwise convey the meaning that the discourse object or set of objects specified by the cleft head is **UNIQUE** with respect to its ability to satisfy the variable contained in the cleft presupposition. Various forms of this claim have been made in the literature. For example, Halvorsen [1978:15] argues that a cleft such as (4-91) creates an **IMPLICATURE** such as (4-92):

(4-91) It was William that dug up the onions.

(4-92) Exactly one individual dug up the onions.

Halliday [1985:43] notes that, in the comparison between a topicalised sentence such as (4-93a) and a reverse *wh*-cleft such as (4-93b) ‘the former implies *among other things*, the latter implies ‘and nothing else’:

(4-93) a A loaf of bread we need.

b A loaf of bread is what we need.

Atlas and Levinson [1981] argue that the ‘uniqueness’ reading is not due to implicature. Instead, they claim that clefts such as (4-94) and (4-95) **ENTAIL** a reading such as (4-96):

(4-94) It was John that Mary kissed.

(4-95) It was John that Mary didn’t kiss.

(4-96) Mary kissed (exactly) one person.

In what follows, we will be looking at the evidence for ‘uniqueness’ in the three types of cleft. We will argue that, while Atlas and Levinson’s explanation of uniqueness appears to deal with the data they present, it is inconsistent with their own account of cleft presupposition. In addition, because of their view that negation is by default external, the account they give of uniqueness in negative clefts is incorrect. Finally, for the same reason, the account in general fails to spot the central role that is played by the presupposition of the cleft in the appearance of a uniqueness reading. We will examine the nature of this role, arguing that a simple and illuminating analogy is available in

the comparison between ‘uniqueness’ as understood for clefts and a similar effect that results from the use of definite referring expressions (discussed at length by Hawkins [1978]). Such an approach not only describes the uniqueness behaviour of clefts simply but falls out of the view of cleft presuppositions as propositions that are interpreted like anaphors, a view which we pursue throughout this chapter.

4.3.1 The Evidence for Uniqueness

It is evident from simple examples that some kind of ‘uniqueness’ reading does exist in clefts. First of all, the cleft in the following context is unacceptable, in spite of the fact that both the entailment *Mary kissed John* and the presupposition *Mary kissed someone* are trivially true in the context:

(4-97) #Mary kissed everyone, and it was John that she kissed.

Modifications of the cleft head that run counter to the ‘unique’ interpretation of its denotation are also unacceptable. For example, the use of *among others* to modify the cleft head creates a non-unique effect that is incompatible with the unique reading, resulting in the unacceptable (4-98):

(4-98) *It was William, among others, who dug up the onions.

Similar effects can be created using *also*. In a context where *also* has scope over the cleft complement, as in (4-99a), the cleft is acceptable. Where it has scope only over the cleft head, which is the phrase denoting the ‘unique’ element, however, the uniqueness reading is violated, resulting in unacceptability. This is shown in (4-99b), below:

- (4-99) Did William dig up the onions?
 a Yes, and it was also William that destroyed the runner BEANS.
 b *Yes, and it was also ZOE that did.

Even has a similar effect. In (4-100a) below, *even* has scope over the cleft complement. As in the previous case, however, where it has scope over the head only, as in (4-100b), it runs counter to the uniqueness reading:

- (4-100) a A: I suppose John helped out with the party, did he?
 B: Yes. It was even John who arranged it.
- b A: I suppose you had a lot of help clearing up after the party, did you?
 B: *Yes. It was even JOHN that helped.

We can conclude, then, that a uniqueness reading certainly seems to exist. But what is uniqueness predicated of? In the introduction to this chapter, we noted that uniqueness was apparently not predicated of single elements. Evidence for this is the fact that plural terms can appear as cleft heads:

(4-101) It was some boys that Mary kissed.

(4-102) It was Mart and Rick that Mary kissed.

In fact, existentially quantified phrases such as proper names, plural and singular definite and indefinite NPs, and generics can all appear freely in cleft head position. Universally quantified expressions, however, are not generally acceptable:⁶

(4-103) *It was every boy that Mary kissed.

How are these facts to be explained? Atlas and Levinson argue that, given an example such as (4-104)

(4-104) It was John that Mary kissed

the uniqueness reading arises from a combination of two factors: the fact that the *it*-cleft carries a presupposition, in this case, *Mary kissed someone*, and the fact that the cleft lists only one person who Mary kissed, namely John. Atlas and Levinson explain the interaction of these two factors as follows:

The sentence *Mary kissed someone* follows from the contribution of *that Mary kissed* to *It was John that Mary kissed*. But the proposition *Mary kissed (exactly) one person* follows because of the CONTINGENT fact that the specification in the (surface) main-clause focus constituent of *It was John*

⁶It is a problem for this account that *wh*-clefts, unlike the other two cleft types, *can* take universally quantified expressions in head position. In the DRT account of presupposition to follow, we will be dealing primarily with *it*-clefts. The distinctive behaviour in *wh*-clefts in this respect suggests that a slightly different treatment would need to be developed for them. We return to this point briefly in chapter 6.

that Mary kissed lists but one item, namely, John. That ‘asserted’ fact adds *Mary kissed (at most) one person* to the ‘presupposed’ *Mary kissed someone* to give *Mary kissed (exactly) one person*.

[Atlas and Levinson 1981:54]

Because this principle is dependent upon the number of items listed by the cleft head, it is possible to extend uniqueness to suggest that Mary kissed more than one person, if more than one person is named in the head.

While Atlas and Levinson’s account seems to be compatible with the data we have seen so far, it is not internally consistent with their view of clefts. It will be recalled from section 4.2 that Atlas and Levinson’s account does not appeal to the notion of presupposition: on their view, positive clefts only *entail* the proposition conveyed by the complement. There is therefore no ‘presupposed *Mary kissed someone*’. If such a view were to be based on entailment, which is a way of making their account consistent, the uniqueness reading should also be associated with normal declaratives. This, however, does not explain the asymmetry between declaratives such as (4-105a) and clefts such as (4-105b) with respect to uniqueness:

- (4-105) a William, among others, dug up the onions.
 b *It was William, among others, who dug up the onions.

While Atlas and Levinson are forced to discover a role for presupposition in the case of the positive cleft to explain the appearance of a uniqueness reading, they argue that no such reading appears in negative clefts; this is in spite of the fact that their analysis of negative clefts postulates the appearance of an implicature which is identical in content to the presupposition required in the positive case.

Apart from this inconsistency, we should also point out that the conclusion that no uniqueness reading exists for negative clefts cannot account for examples such as (4-106):

- (4-106) *Of all the cats I can think of, it wasn’t this cat that dug up the onions.

The explanation behind this unacceptability is as follows. The negative cleft has a uniqueness reading that can be glossed as (4-107):

- (4-107) There exists a unique set of individuals such that they dug up the onions, and this cat is not a member of that set.

However, the function of (4-106) is to postulate *all the cats I can think of* as potential diggers-up of onions. The referent of *this cat* is a member of that set, which violates the uniqueness requirement of the negated cleft, i.e., that the head of the cleft is not a member of the unique set of diggers-up. This suggests that the uniqueness reading is a part of the meaning of the negative cleft in (4-106).

Atlas and Levinson's conclusion that no uniqueness reading exists for negative clefts of this kind may well stem from their view, described in section 4.2, that the default reading of negation is external. If the external reading is taken, the uniqueness reading does indeed disappear. On this reading, we would gloss the meaning of (4-108) as (4-109), rather than as (4-107):

- (4-108) It wasn't this cat that dug up the onions.

- (4-109) It is not the case that there exists a unique set of individuals such that they dug up the onions, and that the sole member of that set is this cat.

In fact, the disappearance of the uniqueness reading is merely a side-effect of the disappearance of the factive presupposition that causes it. In the case of negative clefts interpreted with internal negation, therefore, the existentially-quantified cleft presupposition still serves to assert the existence of a unique set of individuals that can instantiate the variable in the proposition, while the cleft as a whole asserts that the set of individuals currently specified by the cleft head is not a member of that first set. Contrary to Atlas and Levinson's conclusion, there is no reason why the cleft presupposition should lose its force when the negation has scope only over the cleft head.

This view, of course, requires that presupposition plays a central role in both positive and negative clefts, and that the uniqueness reading falls out of the existence of the existentially-quantified presupposition. In support of this view, we can show that this situation is analogous to that occurring in the case of other presupposing constructions, using definite referring expressions as an example.

4.3.2 The Analogy with Definiteness

We would claim the presupposing constructions such as *the* are syntactic signals indicating the existence of a unique set of individuals that satisfy the description accompanying the syntactic signal. In the case of a definite referring expression such as *the bishop*, therefore, the description to be satisfied might be represented as (4-110):

(4-110) $\exists x \text{bishop}(x)$

Analogously, in the case of a cleft such as (4-111), the description to be satisfied is conveyed by the content of the cleft presupposition, which can be represented as (4-112):

(4-111) It was John who kissed Mary

(4-112) $\exists x \text{kissed}(m, x)$

In both cases, an existential presupposition is communicated, which presupposes the existence of a set of elements suitable for satisfying the variable in the expression.

Why, though, should that set be unique with respect to the current context of the discourse? This can be understood in relation to Hawkins' [1978] work on definite referring expressions. Hawkins' claim is that, for the use of a definite referring expression to be appropriate, the expression must refer to the TOTALITY of objects currently in the discourse context that can satisfy the description contained in the referring expression. For example, in the case of the phrase *the bishops*, there must be no bishops in the discourse context that fall outside the set that the speaker intends to indicate by the use of the definite article. If there are, definiteness cannot be used appropriately. We would argue that the cleft presupposition such as that indicated in (4-112b) is analogous: in this case, there must be no objects such that Mary kissed them that fall outside the intended reference of the cleft head—that is, nobody other than John in the current universe of discourse must have kissed Mary if the cleft is to be appropriate. This would explain the unacceptability of the examples we examined at the beginning of this section, such as (4-113). In this example, the element named in the cleft head is only one element of the set of people who Mary kissed, instead of naming the totality of that set:

(4-113) *Mary kissed everyone, and it was John that she kissed.

Finally, we noted in our review of the evidence for uniqueness that both singular and plural expressions could appear as cleft heads—that is, that the cleft head seems to specify SETS of individuals, which may consist of one or more elements. Hawkins' suggestions are compatible with this view, since he proposes that definiteness is predicated of sets. In each case, the set specified indicates the TOTAL or MAXIMAL set of elements to which the description is currently applicable.

In the DRT model we give in the next section, we see how the requirement that the cleft head specifies the MAXIMAL instantiation of the existentially-quantified variable in the cleft's factive presupposition can be implemented as a precondition on well-formed discourse representation structures for dealing with clefts.

4.4 Clefts in Discourse Representation Theory

As we described above, van der Sandt [1988, 1989] has suggested that presuppositions can be analysed as a species of propositional anaphor within the framework of Discourse Representation Theory (henceforward DRT). In this section, we explore this proposal in detail, concentrating for simplicity on *it*-cleft examples. We believe, however, that the treatment we suggest can straightforwardly be applied to *wh*-clefts and reverse *wh*-clefts.

The structure of the section is as follows. We first of all give a brief introduction to DRT. Since we are interested in the treatment of presupposition as anaphora, we will use a simple sentence featuring a pronoun to show how anaphora is treated in DRT. We then go on to look at a propositional presupposition, using as an example the complement of *regret*. Using this example, we demonstrate how the notion of ACCOMMODATION can be dealt with in the framework, before going on to show how a cleft example would be interpreted. In the treatment of the cleft, we draw not only upon the notion of accommodation but also on the discussion of UNIQUENESS, which on our treatment is more appropriately termed MAXIMALITY, and show how these phenomena can be elegantly catered for in the DRT framework. Finally, we examine how the DRT model can account for contexts in which presuppositions fail to survive, capturing this by means of embedded discourse representation structures (DRSS).

4.4.1 A Brief Introduction to DRT

DRT can be described as a formal means of representing discourses and the relationships of the objects that appear within them. Using a representation of discourse which consists of boxes in a variety of relationships, the framework is able to represent constraints on the interpretation of anaphora in terms of the ACCESSIBILITY of potential antecedents. What constitutes an acceptable discourse is defined in terms of constraints on allowable DRSS, using DRS construction rules based on the syntactic structure of the incoming sentences.

In Kamp's original formulation, a DRS for a text is determined by a set of DRS-construction rules which operate top-down on the parse trees of each sentence in the text, thereby using syntactic structure as the basis for the logical representation. Formally defined, a DRS K is a pair $\langle U_K, Con_K \rangle$ where U_K is the *universe* of K , consisting of the set of elements that appear in the discourse, termed the DISCOURSE MARKERS, and Con_K is a set of CONDITIONS consisting of either atomic formulae, or relations on subordinate DRSS. Informally, we can think of the markers as the representations of argument expressions, and the conditions as the representation of predicates or relations between them.

To see how an analysis of a simple NP VP sentence might work, we can consider the example in (4-114):

(4-114) Mary ran.

As we described in the introduction to this chapter, we will be using a notation in which an extra argument position is filled by an EVENTUALITY INDEX. We noted there that, following Bach [1986], eventualities can be divided into the three main categories of STATES, EVENTS and PROCESSES. Examples of each respectively might be the denotations of *be drunk*, *reach the top* and *walk*. Following the notational conventions of UCG (cf. Zeevat *et al.* [1987]), we use a, a_1, a_2, \dots, a_n to range over eventualities, and e, e_1, e_2, \dots, e_n and s, s_1, s_2, \dots, s_n to range over events and states respectively. In a conventional notation, the analysis of the content of (4-114) might therefore be as follows:

(4-115) $\text{ran}(e, m)$

where m is Mary. In DRT, we can draw a DRS K , shown in figure 4.1.

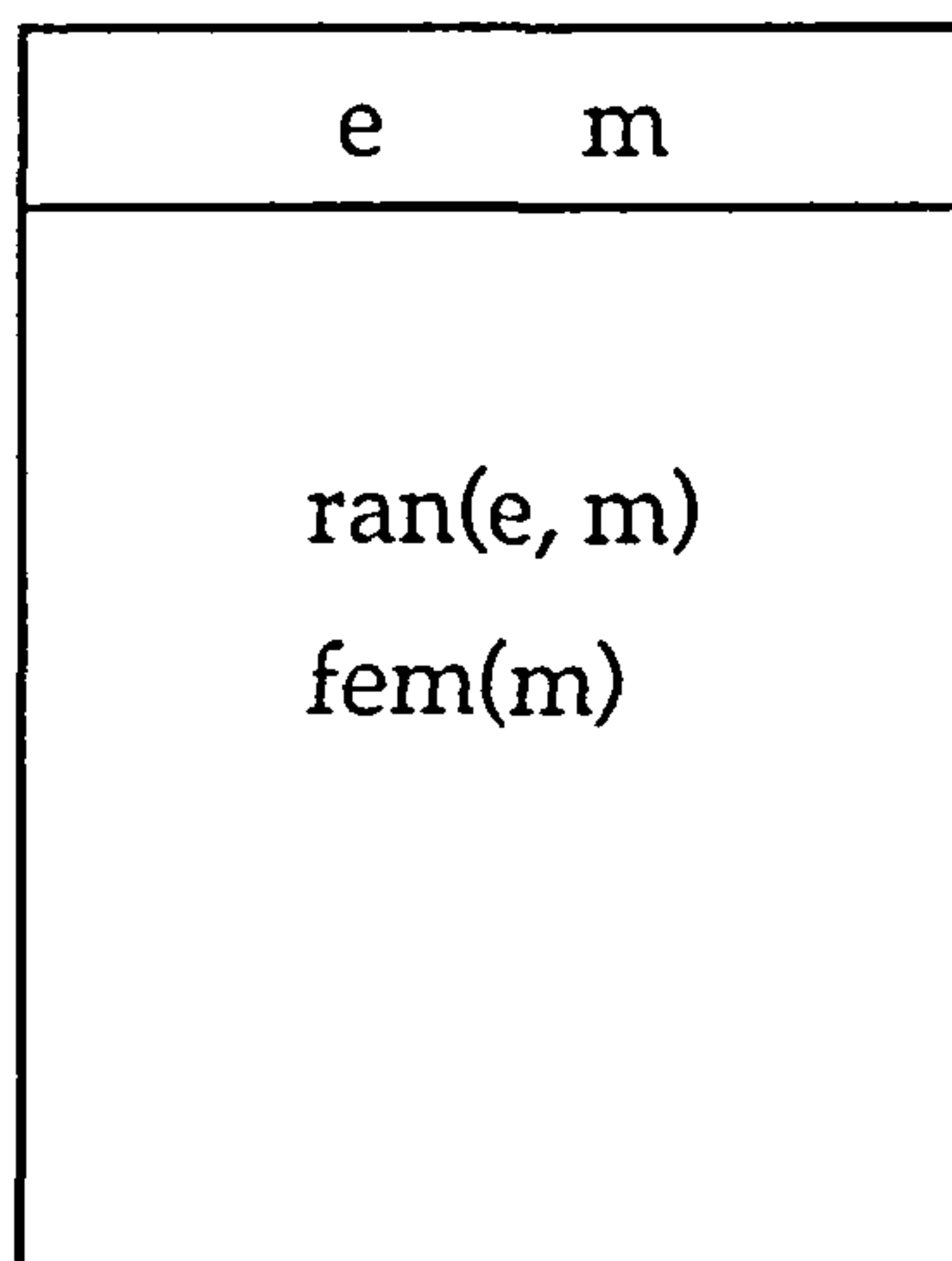


Figure 4.1: DRS for *Mary ran*

To build up a DRS such as that in figure 4.1, we start by analysing the NP, indexing it with a discourse marker m which is inserted into the universe U_K . Since we need to encode the fact that m can act as the antecedent for a feminine gender pronoun, we also add the condition $\text{fem}(m)$.

Next we analyse the VP. This gives rise to an eventuality marker, e , which is inserted into the universe of discourse; it also gives rise to a condition which is added to Con_K , below the line in the box. In many cases, the NP also yields a condition, as we will see in examples below. For reasons internal to the theory, proper names are treated somewhat differently, and do not in isolation introduce conditions.⁷

4.4.2 The Treatment of Pronominal Anaphors

Having seen how a simple, non-anaphoric example would be represented, we can go on to see how pronominal anaphora is dealt with in the theory. The example we will use here is (4-116):

⁷In fact, the NP and VP, when NP is not a proper name, enter into either an conjunctive or an implicative relation, depending on the determiner of the NP. However, we will ignore this issue here, and concentrate only on conjunctive structures.

(4-116) She fell.

In this case, because the NP is the personal pronoun *she*, it gives rise to an instruction to look for a marker that is already present and accessible in the DRS.⁸ In this case, the rule for building the DRS has two parts:⁹

- Find an ‘accessible’ discourse marker x of feminine gender already present in the current DRS.
- Use x as the argument-filling index associated with the NP.

Suppose we take the DRS for *Mary ran* above as our current context; then the result of processing *She fell* in that context is illustrated in figure 4.2.

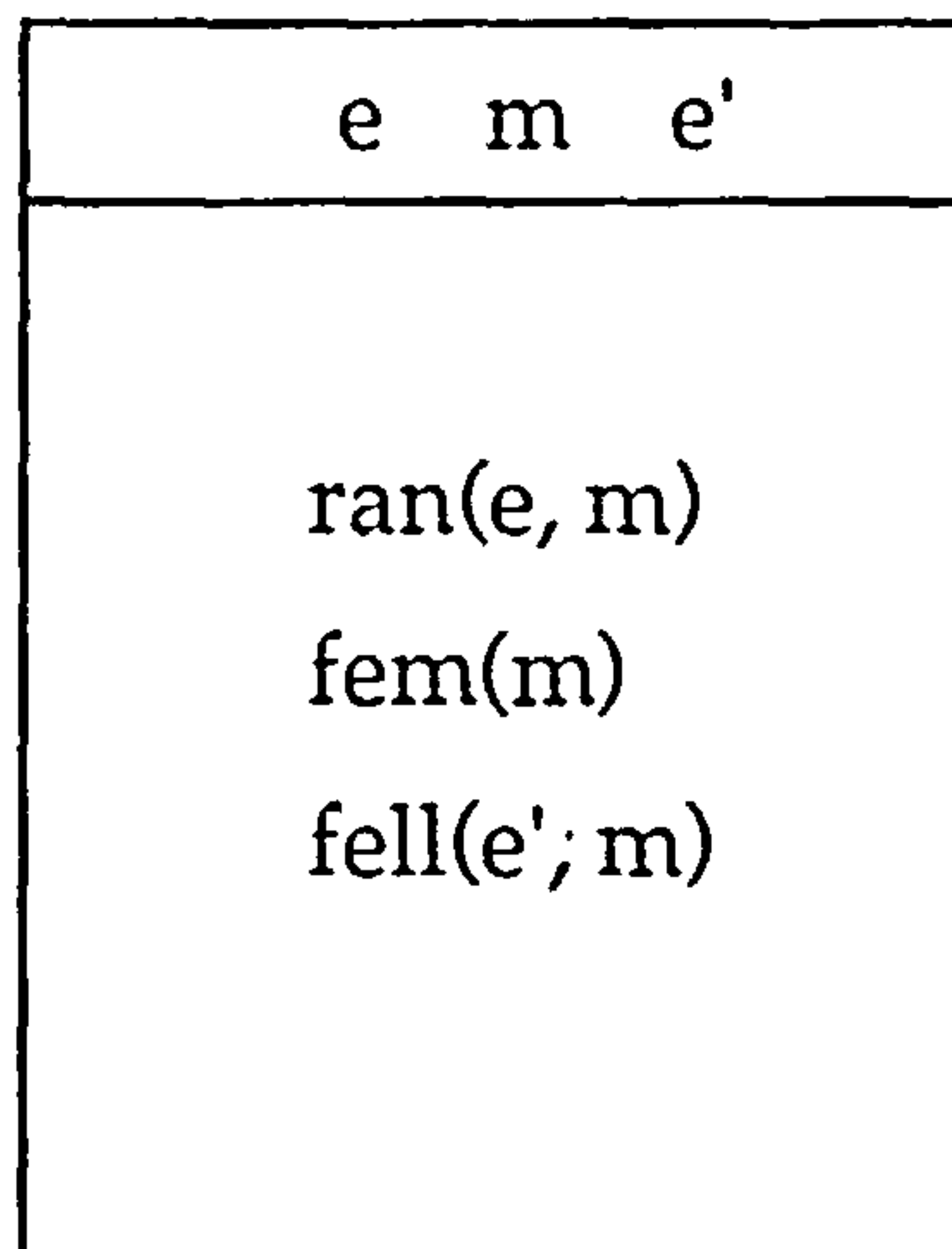


Figure 4.2: DRS for *Mary ran. She fell*

On the basis of this simple example, we will suggest a modification to the representation that will make the treatment of presuppositions clearer and more intuitive. We saw above that van der Sandt views presuppositions and anaphors alike as mechanisms for *context selection*. That is, both place constraints on the contexts in which they are acceptable. We can see these constraints as PRECONDITIONS on what a context—a DRS—has to be like in order for the sentence to be interpreted. In the previous example,

⁸See Kamp [1981] for a precise definition of accessibility. Roughly speaking, the markers which are accessible for a given box K are those in K itself, those in a box which is the antecedent of a conditional where K is the consequent, and those in any boxes which contain K .

⁹We should point out that Kamp [1981] uses a slightly different rule, whereby a new marker y is introduced for the pronoun, and then linked by an equation of the form $y = x$ to the antecedent.

regret. For example, the example in (4-117) presupposes the content of its complement, namely *John ran*:

(4-117) Mary regrets that John ran.

Processing the sentence left to right, we will first of all need to insert a marker for *Mary*. We can then go on to address the VP. Since it includes a stative verb, *regret*, we include a marker *s* indicating the state. The VP itself can be represented as condition consisting of a predicate and three arguments: the first argument, *s*, being the eventuality marker standing for the state introduced by *regret*, the second, *m*, being the argument representing Mary, and the third representing what is regretted, i.e., the presupposition of the sentence. This third argument can be represented as *e*, since it stands not for an atomic element but for another eventuality, the thing being regretted being an action of John's. Because this third argument is also presupposed, however, there is a requirement for the acceptability of the sentence that its content be present in the DRS *prior* to the interpretation of the sentence—that is, its presence is a precondition on the current DRS. We therefore include *e* not only as an argument in a condition on the postcondition side of the construction rule, but as a marker on the precondition side. Finally, we need simply to specify what the content of *e* is. We do this by including its content—*ran (e, j)*—also on the precondition side, since it fills in the content of the presupposed information. The construction rule as a whole appears in figure 4.4.

e j	≡	s m
ran(e, j)		regret(s, m, e)

Figure 4.4: Construction rule for *Mary regrets that John ran*

The analysis just described suggests the following kind of paraphrase for presuppositional constructions featuring *regret*:

- (4-118) a Mary regrets that John ran.
 b There was an event_{*i*} of John running, and Mary regrets it_{*i*}.

Assuming this to be an acceptable paraphrase, it can be observed that an anaphoric relation can be said to exist between the eventuality index *i* in the gloss of the complement of *regret* (the clausal predication *John ran*), and some antecedent proposition in the discourse. In figure 4.4, we represented this as a relationship between the eventuality *e* and its antecedent. However, on the paraphrase in (4-118b), we can formulate the required precondition for presupposition not as the requirement for the presence of the propositional content itself, but as the index of the appropriate eventuality. This move allows us to maintain the generalisation, argued for on independent grounds in Chierchia and Turner [1988], that anaphora is always to arguments, rather than to predicates or relations. It is a standard assumption of UCG (cf. Zeevat *et al.* [1987]) that verbal predicates should be allowed to introduce eventuality indices, on the basis of their similarity with indefinite terms.

For presuppositions to be acceptable, therefore, we are stipulating the following precondition on the DRS into which the presupposition-bearing sentence is to be incorporated: the eventuality index contained in the presupposed clause should already be accessible from the DRS as a marker at the time of interpreting the sentence.

However, as we observed in the discussion of presupposition earlier in this chapter, it is perfectly possible to presuppose information that is not already in the discourse context. As van der Sandt observes, in cases where ‘new’ information is presupposed, this does not lead to an unacceptable discourse. Instead, the presupposition is *accommodated*, to use Lewis’s [1979] terminology. That is, the presupposed proposition is added to the discourse representation as though it were already there. In the terms we are using here, we might say that presuppositions can contain eventuality indices that are not part of the precondition of the DRS at the time of encountering the cleft. The strategy of distinguishing between preconditions and postconditions is intended to be suggestive of the distinction between propositions that are merely entailed, and accommodated presuppositions. In both cases, a proposition is added. But a proposition which is only entailed by a given sentence *S* is added to the DRS as an ‘effect’ of the interpretation of *S*, while a proposition which is an accommodated presupposition of *S* is added as means of satisfying a precondition to the interpretation of *S*.

4.4.4 Cleft Presuppositions as Anaphors

We can now try to integrate our earlier remarks about clefts with van der Sandt's proposal, using the simple *it*-cleft (4-119), the syntax of which can be glossed as (4-120):

(4-119) It was John who ran.

(4-120) It was NP $s[+R]$.

As described in the previous section, the clausal predication contained in the presupposition, in this case $s[+R]$, requires an eventuality index a to act as its antecedent in the DRS. The content of the eventuality—in this case, the event of John running—is also specified as a condition in the precondition, identifying the nature of the eventuality that is required as an antecedent. We also require an additional precondition, because the cleft presupposition is existentially quantified: that is, it contains an existential presupposition that itself requires an additional marker to appear in the DRS.

Having represented the presupposition of the cleft complement as a precondition on the DRS we need to represent the assertion that links the value specified by the head of the cleft to the existentially quantified variable in the presupposition, by means of the cleft copula. We can do this in the manner we described in section 4.2, representing the asserted proposition as follows. In (4-121), s is the eventuality marker introduced by virtue of the stative *be*, while the marker of the referential argument to the copula, x , is anaphoric to the eventuality marker introduced in the context. NP of course represents the value provided by the head of the cleft:

(4-121) $be(s, x, NP)$

This appears as a condition in the postcondition of the representation, since it is not part of the presupposition. The schematic representation of the cleft appears in figure 4.5, while a version with the lexical items filled in is given in figure 4.6.

Note that in the lexically-specified example in figure 4.6, we have been able to specify the kind of eventuality that we are talking about: in figure 4.5, we use the generic eventuality marker a in the precondition; in figure 4.6, we know we are talking about

a	x		s
$S[+R](a, x)$			$be(s, x, NP)$

Figure 4.5: Construction rule for schematic *it*-cleft

e	x		s	j
$ran(e, x)$			$be(s, x, j)$	

Figure 4.6: Construction rule for *It was John who ran*

a running event, and can therefore use the event marker *e* instead. Note further that, while in figure 4.5 we are unable to specify the identity of the denotation of the NP (and cannot include NP as a marker, since NP is not an admissible discourse object) we are able to include in the postcondition that the runner is John; hence the appearance of the extra marker, *j*.

4.4.5 Clefts and Maximality

A final aspect of the meaning of cleft constructions still remains to be incorporated into the representation, however. It will be recalled from section 4.3 that the cleft has a reading associated with it to the effect that the set of discourse objects specified by the

expression appearing as syntactic head of the cleft has to be **UNIQUE** or **MAXIMAL** with respect to the existentially-quantified variable in the presupposed proposition. That is, the cleft head refers to the unique set of elements in the universe of discourse that satisfies the condition expressed by the cleft complement or *wh*-clause. In order to capture the facts we observed, to the effect that the elements specified in this way by the cleft could be singular or plural, we will assume for simplicity that the reference marker X that stands for the head of the *it*-cleft is plural, ranging over sets, and that the apparently singular NPs in preceding examples can be interpreted as singleton sets. We can then account for the exhaustiveness condition along the lines indicated in the revised schema for *it*-clefts given in figure 4.7. In this figure, the presupposition appearing in the precondition is embedded within a further structure, which expresses the requirement that any instantiation of the existentially-quantified variable in the presupposition must satisfy the maximality condition. This condition can itself be expressed as (4-122):

$$(4-122) \quad \text{max}(X, \lambda y \phi) \equiv_{df} \forall x [x \in X \Leftrightarrow \lambda y \phi(x)]$$

In effect, this just says that X is the set $\{x : \phi(x)\}$.

X	s
$\text{max}(X, \lambda x \phi)$	$\text{be}(s, X, \text{NP})$

Figure 4.7: Revised schema for *it*-cleft featuring the maximality condition

Finally, we can see how the schema can be applied to our cleft example in figure 4.8.

Notice that by virtue of (4-122), any DRS which satisfies the precondition of the rule in figure 4.7 will entail $x \in X \Rightarrow \text{run}(e, x)$. Since we have $\{j\} = X$ (by virtue of the postcondition), and therefore $j \in X$, we ensure that $\text{run}(e, j)$ holds in the DRS.

e	X	≡	s
max(X, λx[run(e, x)]		≡	be(s, X {j})

Figure 4.8: Revised DRS for *it was John who ran*

4.4.6 The Projection Problem and DRT

We noted in section 4.2 that a useful model of presupposition would be one that could deal with what is commonly termed the PROJECTION PROBLEM. That is, it would have a way of representing contexts in which presuppositions failed to survive, even though the sentence involved might be one that was usually interpreted as a ‘presupposing’ construction. We also stipulated that the solution to this problem would not consist of adopting the strategy of cancelling presuppositions in these contexts, since, we argued, presuppositions are used intentionally and meaningfully by speakers. A model that cancels the presupposition outright is therefore one that loses part of the intended meaning of the presupposition-bearing sentence. Van der Sandt suggests a way of treating presuppositions in DRT that deals elegantly with the projection problem, by building a structured DRS in which the presupposition is treated as normal, but is subordinated or embedded within the DRS. It therefore fails to be preserved at the top level of the structure, but can be interpreted at the deeper level of structure in an analogous way to that described for presuppositions that do project.

Because we have treated presupposition as a species of anaphora, we require as in other cases of anaphora that an accessible marker exists in the DRS to act as an antecedent for the presupposition. These antecedents can either belong to the current DRS or to one superordinate to it. This is the case even when the presupposition occurs in a context in which presuppositions do not necessarily survive: to say otherwise would mean that the

presupposition could not be interpreted at all, and would be unacceptable. Following van der Sandt [1988], therefore, we will say that when presuppositions are accommodated, regardless of the context in which they appear, they are entered into the highest box that the rules for constructing DRSS permit; i.e., the topmost box that is accessible at the current stage of constructing a DRS. We can say therefore that our preconditions all hold of the topmost admissible box. However, in contexts in which presuppositions may or may not survive, such as conditionals, the content of the presupposition does not appear to be so accessible to subsequent discourse.

We can show how this might take place using an example of a cleft construction that appears as the consequent clause of a conditional. This is represented in figure 4.9.

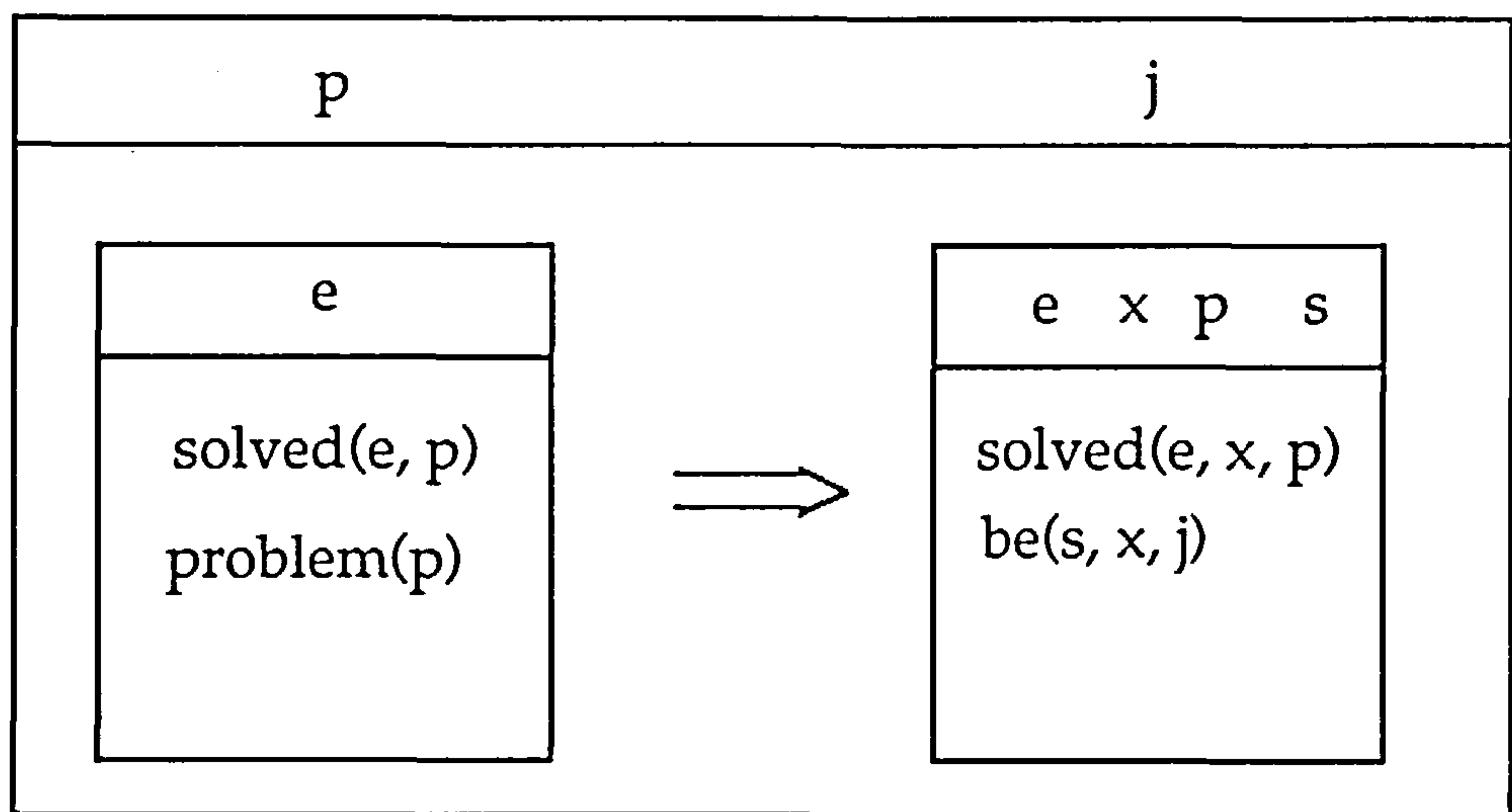


Figure 4.9: *It*-cleft appearing as the consequent of a conditional

Note that in the case of the conditional, an embedded structure is introduced into the DRS. The cleft is analysed as normal, except that the topmost box into which markers (except for those standing for proper names) can be introduced is the embedded box. The presupposition *solved(e, x, p)* therefore does not hold of the superordinate box of the DRS.

We would claim that the DRT treatment of cleft presupposition proposed is an elegant and intuitively appealing way of dealing with the semantic aspects of clefts we have examined in this chapter. Before going on to look at how our conclusions here can address the problems presented by a corpus of clefts, however, we will summarise the

conclusions we have arrived at so far.

4.5 Aspects of Cleft Semantics: Summary

In this chapter, we hope to have brought out some important aspects of the meaning of clefts. First of all, we have claimed that cleft presupposition is a syntactic means of signalling to a hearer or reader that some antecedent should be found in the discourse context for the presupposed material. At the same time, the cleft conveys an assertion that specifies a the maximal set of individuals in the discourse context that can act as a value that can instantiate the existentially-quantified variable in the presupposition of the cleft. We argued that an appropriate treatment of these aspects of the meaning of the cleft, unavailable in other theories of presupposition, could be found in the DRT model of presupposition suggested by van der Sandt [1988, 1989]. Pursuing van der Sandt's suggestions, we developed a model for cleft presupposition, and showed how the absence of appropriate antecedents for cleft presuppositions could be dealt with through an accommodated eventuality index for the presupposed material. Finally, we demonstrated how DRT provides a mechanism that could capture the linguistic behaviour of clefts in contexts where presuppositions fail to project, without losing the information that a presupposition has meaningfully taken place in the discourse.

In the next chapter, we will see how the phenomena and processes we describe in this chapter can be applied in the description of a corpus of data. While we have explored the nature of some major semantic and pragmatic factors that are relevant in the description of cleft constructions in general, we concentrate in the next chapter on the examination of exactly how, and to some extent why, these factors come into play in the use of clefts in context.

Chapter 5

Pragmatic Aspects of Cleft Constructions

In chapter 4, we concentrated on how part of the interaction between clefts and context, namely the function of clefts as PRESUPPOSITIONAL constructions, could be captured in formal terms. In this chapter, we extend that view in two ways. First, we attempt to capture the variety that exists in the relationship between cleft presuppositions and context. Secondly, we examine other aspects of the coherence relations of clefts as a whole with the surrounding context, employing some of the same formal machinery. The aim of the chapter as a whole is to capture the range of relations that can exist between the content of a cleft and the context, specifying which aspects of these relationships are obligatory, and which are subject to variation.

In fulfilling this aim, we examine four separate but related aspects of clefts in discourse, as follows:

- the accentual patterns occurring in the data;
- the factors surrounding the serial ordering of information that are relevant to the use of clefts;
- the precise relationships that exist between cleft presuppositions and context across a range of data; and
- the relationships that exist between the content of clefts as a whole and the context.

Several of these areas have already been touched upon in the discussion of previous research in chapter 2: part of the aim here, therefore, is to further substantiate the claims we made there, and show in particular how prevalent the counterexamples we produced actually are in a corpus of data. Our main aim in the discussion of each area, however, will be to point out the distinctions that exist between each of the three cleft types, underlining how a choice between them might be made in a given context.

We deal with the four issues listed above in sections 5.1–5.4 respectively. In section 5.5, we present a synthesis of the relative roles of syntactically-signalled presupposition on the one hand, and information structure on the other, in the production and interpretation of clefts. We also relate our conclusions to the psychological model of discourse processing suggested by Clark and Haviland [1977]. Finally, in section 5.6, we give a summary of the claims made in the chapter with respect to the functions of clefts in discourse.

The Corpora

Two sub-corpora, which we will refer to as sub-corpus A and sub-corpus B, are used in the analysis presented in this chapter. Sub-corpus A consists of 150 examples, 50 of each of the three cleft types, and is taken at random from the spoken corpus of clefts derived from the Survey of English Usage (cf. Svartvik and Quirk [1980]).

As we noted in chapter 1, the survey provides detailed information concerning the prosody of the spoken data, and sub-corpus A is accordingly the one upon which we base our discussion of the prosodic features of the clefts.

Sub-corpus B was chosen from the corpus of spoken data as a whole, including my own collection of data. In this corpus, prosodic information was not required: it was simply preferable to draw the data from the most comprehensive source available. Sub-corpus B is therefore partially overlapping with sub-corpus A. It too contains 150 examples, 50 of each type of cleft.

The corpora used throughout this thesis appear in the appendix. Those examples which constitute sub-corpora A and B are distinguished by the appropriate letter appearing

next to the examples in the main corpus.

The Notation

In the discussion, we will be using a linearised version of the DRT notation adopted in chapter 4. The markers currently in the universe of discourse will be bounded by square brackets, while conditions will be represented the same as before, except in linear form. For example, the representation corresponding to (5-1) is shown as (5-2):

(5-1) John ran.

(5-2) [e, j] ran(e, j)

In the discussion of prosodic patterns, lexical items upon which nuclei appear are indicated in small capitals, while those on which subsidiary accents appear are italicised. For example, the cleft in (5-3) features a single nucleus on the head element, and a subsidiary accent in the complement:

(5-3) It's LIBERTY who's making all the *noise*.

5.1 The Accentual Patterns of Clefts

In chapter 2, several claims with respect to the accentual patterns of clefts were examined. We saw that several prevalent assumptions about the possible accentual patterns rested on too narrow a view of the data, or were based on an ill-founded expectation that accentual pattern would correlate reliably with syntactic structure. Some particular assumptions discussed were as follows:

- the assumption that the head constituent of a cleft always bears a nuclear accent (cf. Chafe [1976], Creider[1979]);
- the assumption that both the head and the complement of the cleft bear a nuclear accent (cf. Quirk *et al.* [1985]); and
- the assumption that clefts consist of a single tone group (cf. Quirk *et al.* [1985]).

In chapter 4, during the course of the discussion of theories of presupposition based on sentence accent, we reviewed the following further assumptions:

- the assumption that clefts contain a single nuclear accent (cf. Chomsky [1971]); and
- the assumption that *wh*-clefts do not contain nuclei in the *wh*-clause (cf. Prince [1986]).

We examined counterexamples to each of the above at the time, and therefore do not intend to review the arguments here. Instead, the purpose of this section is simply to describe the five major accentual patterns of clefts, and expose the prevalence of each in a small corpus of data. In addition, the composition of clefts of each type in terms of the number of tone groups each consists of on average will be discussed briefly. In the analysis, we assume that CONTRASTIVE accents can be either subsidiary accents or nuclei; they are not considered to be part of a separate system.¹

5.1.1 Five Types of Accentual Pattern

Five kinds of accentual pattern were observed in the data, as follows.

Head-Nuclear Clefts in which the cleft head has a nuclear accent, and no accent appears in the cleft complement;

Complement-Nuclear Clefts in which the cleft complement or *wh*-clause has a nuclear accent, and no accent appears in the cleft head;

Both-Nuclear Clefts in which a nucleus appears in both the head and the complement of the sentence;

Head-Nuclear/Weak Complement Clefts in which a nucleus appears in the cleft head, and a subsidiary accent occurs in the cleft complement; and

Complement-Nuclear/Weak Head Clefts in which a nucleus appears in the cleft complement, and a subsidiary accent occurs in the cleft head.

¹This is the position taken by Gussenhoven [1983] in his research into accent placement, which considers contrast to be an instance of an accent of any kind having 'narrow scope' over the information conveyed by the utterance. Others taking this view include Ladd [1983].

Examples of each type appear below.

Head-Nuclear Clefts

The HEAD-NUCLEAR pattern appears in (5-4) and (5-5):

(5-4) it was a very educative PERIOD that I spent there

(5-5) B: but I have no reason to believe he teaches linguistics

A: no

B: you see or that they have any reason for wanting a teacher of linguistics
what he teaches I think is modern LANGUAGES

No reverse *wh*-clefts with this accentual pattern appear in the 150 sentences examined in this part of the study, but they do appear in my own collection of spoken data. These cases are ones in which a nucleus appears on the cleft head element to indicate contrast, but the head is anaphoric, referring to Old information:

(5-6) JD: That was my friend Mike.

MM: Gosh. THAT's what he looks like.

Complement-Nuclear Clefts

The second pattern is the COMPLEMENT-NUCLEAR pattern, exemplified in (5-7) and (5-8):

(5-7) B: Joe Wright you mean

A: yes yes

B: I thought it was Joe Wright who'd walked in at FIRST

(5-8) B: we're big enough to stand on our own feet now and this is what
Vincent said NO about

No *wh*-cleft appeared either in sub-corpus A or in the corpus as a whole with the complement-nuclear pattern: in *wh*-clefts, any nucleus appearing in the *wh*-clause is counterbalanced by at least one nucleus on the head constituent.

Both-Nuclear Clefts

The third pattern, BOTH-NUCLEAR, is as its name suggests one in which a nucleus appears in both the head and the complement of the cleft. This pattern is compatible with all three kinds of cleft, as demonstrated in the examples below:

- (5-9) A: now where did I hear that from
B: probably me on the phone was it it was the day AFTER when I RANG
- (5-10) A: but have you got a kettle
B: well what I would USE is one of those little SOLID FUEL jobs—they're awfully good and they're quicker than the kettle
- (5-11) A: 'it's perfectly clear from the notes'
B: THIS is what's so AWFUL about academics isn't it—this is the worst side of them

80% of *wh*-clefts and 50% of *it*-clefts in the sample are of this pattern.

Head-Nuclear, Weak Complement Clefts

Examples of the Head-nuclear/weak complement type are as follows:

- (5-12) B: it's the grammar where the fun is
A: yes it's the GRAMMAR which is *interesting*
- (5-13) A: surely they're all happy
B: mm
A: no what they're *having* is the school PLAY you see

No examples of reverse *wh*-clefts of this kind appear in sub-corpus A. They do, however, appear elsewhere in the corpus: the following example is taken from the segment immediately succeeding our sample (reference 2.6.664 in the SEU corpus).

- (5-14) C: He wants to be near the British Museum to go in for Spenser Variorum
—do you think THAT's what it is

Type	<i>it</i> -cleft	<i>wh</i> -cleft	Reverse <i>wh</i> -cleft
Head-nuclear	5	4	0
Complement-nuclear	13	0	43
Both-nuclear	25	40	4
Head-nuclear/weak complement	4	6	0
Complement-nuclear/weak head	3	0	3

Figure 5.1: Breakdown of the sample into the five accentual patterns

Complement-Nuclear, Weak Head Clefts

The complement-nuclear/weak head pattern appears in *it*-clefts and reverse *wh*-clefts, but not in *wh*-clefts. *It*-cleft and reverse *wh*-cleft examples appear below:

- (5-15) A: Yes but you see it doesn't really make any difference you see what they've got it's how much they *move* it that COUNTS
- (5-16) A: I said I wanted some coaxial cable as what I've been using on the old TV thing 'cos I was suspicious - I've still got a hope that *that's* what's WRONG because you know I was in when we got this house that television cable

The absence of *wh*-clefts of this type is due to the same regularity encountered in the case of complement-nuclear *wh*-clefts: nuclei in the complement do not appear in the *wh*-clefts in the corpus unless balanced by a nucleus in the *wh*-cleft head.

Having looked at examples of the five accentual patterns, we can look at their distribution across sub-corpus A, shown in figure 5.1. No statistical conclusions should be drawn from this distribution, since the population in several of the cells is too small; it may be suggestive of tendencies, however.

Note that the accentual patterns we observed in chapter 2 to be the assumed default pattern for clefts of all kinds, namely the HEAD-NUCLEAR pattern with no accents elsewhere in the cleft, has the second smallest representation in the data: 10% of *it*-clefts, 8% of *wh*-clefts, and no reverse *wh*-clefts. From this data, in addition, we can say that reverse *wh*-clefts are typically complement-nuclear, *wh*-clefts are typically both-nuclear, and *it*-clefts are typically either both-nuclear or complement-nuclear.

5.1.2 Profile of Tone Groups in the Data

A tone group is an abstract phonological unit which is used to describe the phonological characteristics of stretches of speech longer than syllables, having a recognisable internal structure and fairly characteristic properties (although definitions differ: for an overview of this and similar notions, see Couper-Kuhlen [1986:73ff]). The data in the SEU corpus (and therefore in sub-corpus A, since it is a subset of the SEU data) is analysed into tone groups on the basis of the definition proposed by Crystal [1969:205]. Crystal's tone unit is defined according to the presence of a peak of prominence in the form of a nuclear pitch movement (the nucleus), followed by a boundary indicated by a perceivable pitch change, the direction of which depends upon the direction of movement on the nucleus. In addition, boundaries may be indicated by other junctural features such as pause and variations of length and aspiration. All this will be of interest mostly to the analyst interested in pursuing research on the basis of the results given here; most relevant to note in general is simply that, given the definition of tone group as having a nuclear accent, and the assumption that nuclei correlate with New information, the number of tone groups in utterances permits a crude reckoning of the 'informativeness' of the utterances involved. The data on tone groups are therefore interesting in that they do give a rough approximation of how relatively informative the cleft types are in use. However, for various reasons we will discuss in relation to the study of information structure in section 5.4, in this thesis we will not be adopting an approach that relies on phonological data for the investigation of information structure.

A count of the number of tone groups in each cleft of the sample appears, according to syntactic type, in figures 5.2, 5.3, and 5.4, for *it*-clefts, *wh*-clefts, and reverse *wh*-clefts respectively. The range of values was as follows: *it*-clefts were found to consist of between 1 and 6 tone groups, *wh*-clefts of between 1 and 10, and reverse *wh*-clefts of between 1 and 3 tone groups. The average number of tone groups for each group was as follows: *it*-clefts, 1.68; *wh*-clefts, 6.48; and reverse *wh*-clefts, 1.34.

If this analysis is to be taken as indicative of general tendencies, on the assumptions regarding the correlation of tone units with New information the conclusion would be that reverse *wh*-clefts contain the least New information of the three cleft types, whereas *wh*-clefts contain the most.

No. of tone units	1	2	3	4	5	6
Frequency	30	11	5	2	1	1

Figure 5.2: *It*-cleft frequencies across tone group counts

No. of tone units	1	2	3	4	5	6	7	8	9	10
Frequency	5	12	13	9	5	3	2	0	0	1

Figure 5.3: *Wh*-cleft frequencies across tone group counts

5.2 Clefts and the Serial Ordering of Information

In several of the accounts of cleft function discussed in chapter 2, we encountered the hypothesis that cleft constructions arise from a need to present a message in some particular serial order—be it SVO order (cf. Geluykens [1984]), or various versions of the ‘Given Information before New Information’ hypothesis (also cf. Geluykens [1984] and Werth [1984]). In that chapter, we saw that these claims were not completely supported by the data, since several cases could be found in which a cleft had been used, but the ordering of the information in the cleft sentence was no better on the salient criteria of the theory in question than the corresponding canonical declarative sentence would have been. Although these serial ordering hypotheses had to be discounted as explanations for the existence of clefts, however, this is not to say that serial ordering principles are irrelevant to the discussion of cleft constructions. In particular, they appear to be highly relevant when it comes to the decision as to which of the three types of cleft sentence will be appropriate in a given context.

In this section, therefore, we will take a brief look at the evidence for an ordering principle that is more general than the problematic ‘Given before New’ principle discussed

No. of tone units	1	2	3
Frequency	41	6	3

Figure 5.4: Reverse *wh*-cleft frequencies across tone group counts

in chapter 2, but in some cases makes the same predictions. The principle is based on the notion of REFERENTIAL CONTINUITY culled from the psycholinguistics literature (cf. Garnham, Oakhill, and Johnson-Laird [1982], Johnson-Laird [1983:372ff]). Parallels between this research and work in linguistics in the field of THEMATIC CONTINUITY are also relevant, and will be pointed out where appropriate.

The Notion of Referential Continuity

The importance of referential continuity for clefts hinges on the issue of the ease of comprehension of sentences, and the assumption that arrangements of information that are easy for hearers to comprehend will be preferred by speakers over other information arrangements. Garnham, Oakhill, and Johnson-Laird [1982] demonstrated experimentally that randomized stories were much harder to interpret and remember than stories where the continuity of devices linking sentences to one another was uninterrupted. This effect is perhaps best illustrated in terms of some simple examples along the lines of the stimuli constructed by Ehrlich and Johnson-Laird [1982] who conducted a follow-up experiment to test Garnham *et al.*'s conclusions. They found that descriptions of a simple scene, a table place-setting in these examples, were much easier to process if the description allowed the subject's model of the scene to be built up continuously. (5-17) is an example of such a continuous description:

- (5-17) The knife is in front of the spoon.
The spoon is on the left of the glass.
The glass is behind the dish.

In (5-17,) the order of reference to the objects involved is *knife-spoon-spoon-glass-glass-dish*. The following example, in contrast, exhibits the referentially discontinuous progression *glass-dish-knife-spoon-spoon-glass*, which was predicted to be much harder to process:

- (5-18) The glass is behind the dish.
The knife is in front of the spoon.
The spoon is on the left of the glass.

As predicted, subjects took significantly longer to process and integrate the final statement of discontinuous examples such as (5-18) than was the case for referentially con-

tinuous examples, and a mean of a second longer to process the description as a whole.

As we noted above, the principle that referential continuity should be preserved makes the same predictions in some cases as the Given-before-New principle. For example, in (5-17), *spoon* is mentioned before *glass* in the second statement, which amounts to the placement of Given before New information. However, there is more to accounting for the coherence of (5-17) than simply this. In the third statement of (5-18), for example, both *spoon* and *glass* have been mentioned before, and so are both Given. The Given-New distinction is therefore unable to predict which of the two NPs should appear first. The preservation of referential continuity refers therefore not just to Given-before-New, although this may be a side-effect of it.

Ehrlich and Johnson-Laird's conclusions, however, also suggest a parallel with an area that has long been the concern of text linguists, the study of THEME and THEMATIC PROGRESSION. Much research has been directed at this issue, resulting in several well-known characterisations of the typical thematic progressions in various languages (see, for example, Firbas [1964, 1966, 1975], Enkvist [1973], and Daneš [1974]). While it is not to the point here to go into this area in detail, it is useful to point out the parallel between this and psychological work on the processing of texts in order to clarify the point that work on theme and thematic progression, and work on integrating propositions into memory, are in fact different approaches to the same problem. For example, the thematic progression in (5-19) is an example of Daneš' [1974] LINEAR PROGRESSION, the thematic structure that in Ehrlich and Johnson-Laird's terms observes referential continuity most closely:

- (5-19) William has just eaten a crane-fly.
It flew in through the window.

This example, although less complex than (5-17), demonstrates referential continuity in the same way, since the two references to the crane-fly—*a crane-fly* and *it*—are maximally close together, uninterrupted by intervening material. In experiments reported by Glatt [1982], conducted along similar lines to those reported by Ehrlich and Johnson-Laird, the two kinds of linear progressions defined by Daneš [1974] were found to be easier to process than the other progressions tested. Note that (5-19), as in many of these examples, also features an anaphoric relation: an effect of the preservation of refer-

ential continuity is often that anaphoric elements remain near their antecedents. The preservation of proximity in anaphoric links also appears to be a priority in discourse: Hobbs [1978] and Pinkal [1986:370] have noted that the vast majority of pronouns, for example, relate to an antecedent that is either in the same sentence, or the immediately preceding one, which indicates that such proximity is desirable.²

Referential Continuity and Cleft Type

The relevance of these points to the use of clefts in discourse can be illustrated by means of a general principle and some examples. The general principle is that, if referentially continuous texts and discourses are easier to process, it is likely that the co-operative speaker or writer will attempt to construct texts and discourses that preserve referential continuity, and the proximity of anaphors to their antecedents, wherever possible. The relevant examples to point out are those in which the preservation of referential continuity as a principle is able to decide between rival sentence types for a given context. One such example, drawn from the corpus of written data, appears below:

- (5-20) Part of Magritte's school holidays was spent with grandmother and an aunt at the town of Soignies. *It was there, whilst playing one day with a small girl in an abandoned cemetery, that he first became aware of painting as a special and somehow mysterious activity:* 'We used to lift up the iron gates and go down into the underground vaults. Regaining the light again, I found, in the middle of some broken stone columns and heaped-up leaves, a painter who had come from the capital, and who seemed to me to be performing magic.'

In this example, the referential chain is, in outline, *the town of Soignies—there—painting—the painter*. The substitution of a *wh*-cleft in the same text is much less acceptable, since it leads to the chain *the town of Soignies—painting—there—the painter*, where the referential dependencies between the two expressions are crossed:

²For a discussion of how anaphors are processed, see for example Sanford and Garrod [1981].

- (5-21) Part of Magritte's school holidays was spent with grandmother and an aunt at the town of Soignies. *Where he first became aware of painting as a strange and mysterious activity, while playing one day with a small girl in an abandoned churchyard, was there:* 'We used to lift up the iron gates and go down into the underground vaults. Regaining the light again, I found, in the middle of some broken stone columns and heaped-up leaves, a painter who had come from the capital, and who seemed to me to be performing magic.'

Finally, the chain of reference that would be produced by a declarative in this context is as discontinuous as that of the *wh*-cleft in (5-21); the sentence itself is more acceptable than the *wh*-cleft, although less preferable than the *it*-cleft. The chain produced by a declarative would be *the town of Soignies—painting—there—the painter*.

- (5-22) Part of Magritte's school holidays was spent with grandmother and an aunt at the town of Soignies. *He first became aware of painting as a strange and mysterious activity while playing one day with a small girl in an abandoned churchyard there:* 'We used to lift up the iron gates and go down into the underground vaults. Regaining the light again, I found, in the middle of some broken stone columns and heaped-up leaves, a painter who had come from the capital, and who seemed to me to be performing magic.'

On the basis of this evidence, we can say that serial ordering is likely to be important in the choice between the three cleft types. However, we do not suggest this as the sole motivation for using a cleft, and we do not motivate the notion of serial order in terms of Given and New information, or *svo* order.

We will briefly touch upon this and a related processing constraint in the discussion of how cleft types tend to correlate with particular presuppositional relations, in section 5.3. In addition, suggestions are made in chapter 6 for how a principle of referential continuity could be incorporated into a decision procedure for choosing between clefts.

5.3 Cleft Presuppositions in Discourse

In chapter 2, we reviewed some claims with respect to the discourse status of the presuppositions borne by cleft complements. We noted in particular a claim made by Prince [1978] to the effect that *wh*-clefts presuppose GIVEN information, while *it*-clefts presuppose information that is KNOWN, defined respectively as follows (Prince [1978:903,

Given information: Information which the co-operative speaker may assume is appropriately in the hearer's consciousness.

Known information: [Information that is] *not* marked as assumed to be in the hearer's mind ...often, though not always, known from the context ...information which the speaker presents as being factual and as already known to certain persons (often not including the hearer).

It was argued in response to these claims that such a qualitative distinction between the presupposed information in each case could not be upheld, on two grounds. Firstly, we saw that *it*-clefts and *wh*-clefts were interchangeable in certain contexts, which suggested that the presupposed information they conveyed could not *always* be different in the way Prince suggests. Secondly, the presuppositions of *it*-clefts spanned such a wide variety of information statuses—from shared, highly salient information, right through to information previously unknown to the hearer, and not mentioned in previous discourse—that it was unlikely that any space on the presuppositional spectrum remained that could be called the exclusive domain of *wh*-clefts. On this basis, it was argued that *wh*-clefts and *it*-clefts inhabit the same CONTINUUM of presupposition types, the distinctions between them arising from their occupying partially distinct, but overlapping, domains.

In this section, we examine the range of presupposition types that appear regularly in clefts. First of all, we look at a particular rhetorical constraint that seems to preclude propositions rhetorically related to previous discourse in that way from being presupposed at all. Secondly, we go on to look at what presuppositional relations can and do occur, and establish a formal means of looking at them, as follows.

In the DRT model developed in chapter 4, we argued that, for a presupposition to be used acceptably in discourse, the presupposed proposition requires an ANTECEDENT in the discourse context in the form of a condition representing the presupposed material, and a suitable eventuality index. We drew a distinction between presupposed propositions whose antecedents were present in the discourse context at the time of processing the cleft construction, and those whose antecedent propositions were not already present. In the latter case, we argued that the antecedent to the presupposed material had to be

constructed in context—ACCOMMODATED, following Lewis [1979]—before the content of the cleft sentence as a whole could be interpreted. In this section, we postulate three types of relation that can exist between a presupposition and its antecedent in context: presuppositions can be OLD, INFERRABLE from context, and NEW. These categories may also be divided into various subtypes, some of which have been identified before with respect to *wh*-clefts, by Prince [1978] and Geluykens [1984]. The similarities between the approach taken here and those of Prince and Geluykens are noted when they occur. The categories suggested here apply in general to all three types of cleft, representing a broader range of data than that described by Prince and Geluykens.

After identifying the statuses of the presuppositions involved, we go on to look at the distribution of the cleft types found in the data with respect to these statuses. We conclude that there are important distinctions between the cleft types with respect to the presupposition statuses generally employed, the most categorical of these being that *wh*-clefts cannot carry presuppositions that appear at the NEW end of the range. Some of these regularities are explained at the time; others, being due to more general principles, are deferred until the appropriate part of the exposition. The framework we develop in this section for looking at presuppositional relations is also used in section 5.4 as part of the mechanism for determining the INFORMATION STRUCTURE of clefts.

Having described the aims and function of this section, we can go on to look at the first issue relevant in the description of cleft presuppositions: a rhetorical constraint on the kinds of proposition that can be presupposed.

5.3.1 A Rhetorical Constraint on Presupposition

Notably absent from the set of relationships observed between cleft presuppositions and the preceding discourse context is a rhetorical relationship that can be characterised as SIMPLE NARRATIVE SUBSEQUENCE. That is, clefts in general do not seem to be used for the simple purpose of relating two subsequent events in a narrative. This can be demonstrated by the following set of examples. (5-23) is a constructed discourse consisting of two declarative sentences, each of which describe an event. The two events, each referred to in the simple past tense, are assumed to be temporally serial:

- (5-23) Mary left the house.
She went to the shops.

When the same content is conveyed with an *it*-cleft³ in place of the second sentence, the resulting sentence is unacceptable:

- (5-24) Mary left the house.
#It was she who went to the shops.

It is only possible to construct an acceptable discourse where the second sentence is a cleft, apparently, when the cleft presupposition conveys information about Mary that can be interpreted as not temporally serial in the discourse—‘background’ information⁴ about her, for example, in the sense that it does not contribute to the main narrative sequence of the discourse. Such a discourse can be achieved, for example, by changing the tense of the verb in the complement to pluperfect, thereby destroying the relationship of narrative subsequence between the two events described:

- (5-25) Mary left the house.
It was she who had been to the shops once already.

Narrative subsequence is also destroyed if the eventuality described by the second sentence is not one which would plausibly be assumed, in the absence of explicit temporal indicators, to be a temporal successor to that described in the previous sentence. The information that Mary had particular feelings about the shops, for example, would be an example of such an eventuality. The resulting sentence, predictably, can be a cleft:

- (5-26) Mary left the house. It was she who loved the shops.

This evidence, combined with the set of relationships that were found to occur between cleft presuppositions and the preceding discourse, suggests that narrative subsequence is not a link that can be used acceptably by cleft presuppositions.

It is likely that other rhetorical relationships exist that similarly preclude the presupposing of the proposition in question. As we suggest in chapter 6, further research of

³An *it*-cleft is chosen for this example because, as will become clear below, *it*-clefts are the type of cleft able to convey the broadest range of presupposition types. Similar examples can, however, be constructed using *wh*-clefts and reverse *wh*-clefts.

⁴It is not intended here that the term ‘background’ be interpreted as meaning ‘less salient in the discourse’. As we noted in chapter 2, the information conveyed by cleft complements is equally salient in the discourse as information carried elsewhere in the cleft when measured in terms of the frequency of subsequent references in discourse.

a comparative nature on a corpus of both cleft and non-cleft sentences would be necessary to establish the nature of these relationships, and to assess how far any general principles can be adduced to explain them.

Having noted this apparently general constraint on what can be presupposed, we can go on to the characterisation of the kinds of presuppositional status observed in the clefts in sub-corpus B.

5.3.2 Three Types of Presuppositional Relation

Sub-corpus B yielded cleft presuppositions of three statuses, as follows:

Old: where a proposition that arises directly from an explicit utterance matches the presupposed proposition;

Inferrable: where the cleft presupposition has no antecedent proposition in the discourse context, but one can be created for it by inference; and

New: where the cleft presupposition has no antecedent proposition in the discourse context, and none can be created for it by inference.

Of the three presupposition types, the latter two require the content of the presupposed proposition to be accommodated in the discourse context. These two types can be seen as requiring or specifying two separate strategies by which accommodation can be performed. The three types are described and exemplified below.

Old Presuppositions

As noted above, a cleft presupposition is Old when it is entailed by the discourse context at the time of processing the cleft. We can point to two subtypes of Old proposition: in the first type, which we will refer to simply as a case of **IDENTITY** between the presupposed proposition and a proposition in context, the presupposed proposition is straightforwardly entailed by the context. For example, (5-27) contains a constructed

context S1 followed by a cleft S2⁵ where the cleft presupposition is entailed by the context:

- (5-27) S1: Who cooked the carrots?
S2: It was Mary who cooked the carrots.

In the linear DRT notation described in chapter 4, we might represent S1 as the set of condition C1, as follows:

- (5-28) C1: $[e_1, x, a]$ cooked(e_1, x, a) \wedge carrots(a)

We can represent S2 in a similar notation, as conditions C2 (the presupposition) and C3 (the assertion):

- (5-29) C2: $[e_2, z, b]$ cooked(e_2, z, b) \wedge carrots(b)
C3: $[s, m]$ be(s, z, m)

In this case, the presupposition C2 is a DIRECT MATCH with C1: that is, C1 is currently available as a propositional antecedent in context. We can say in these cases that a proposition already exists in context that can UNIFY with the cleft presupposition.

We also come across cases where a proposition that describes an eventuality of the same TYPE appears to be acting as an identical propositional antecedent. In this second type of Old presupposition, which we will term TYPE IDENTITY, the cleft presupposition does not strictly have a direct match in the context, since the eventuality described by the proposition that is treated as does not match in every detail. Motivation does exist, however, for treating these examples as special cases of Old presupposition, in terms of a similar phenomenon that occurs in VP ellipsis. The anaphoric relation to which we refer is termed SLOPPY IDENTITY, and occurs when an ellipsis has as antecedent content that is non-identical to the elided material, but which is of the same type. This is exemplified in (5-30), the elided material being indicated by the symbol ϕ :

- (5-30) John made a cake, and Mary did ϕ too.

In (5-30), the cake that John made is not the same cake that Mary made. This case is exactly analogous to cleft cases such as (5-31):

⁵Note that in S2 it would be more natural to omit the complement of the cleft. It is allowed to remain in this case in the interests of clarity.

- (5-31) S1: On Tuesday, John made a cake.
 S2: On Wednesday, it was Mary who made a cake.

In (5-31), two non-identical cake-making events are taking place, but they are treated as identical for all intents and purposes. Evidence for this is that the *it*-cleft complement can be omitted in these cases:

- (5-32) S1: On Tuesday, John made a cake.
 S2: On Wednesday, it was Mary ϕ .

For a full formal treatment, such cases require a notation that allows generalisation over types of events. For simplicity, we can represent these examples as having different eventuality indices. In the following example, C1 represents S1; C2 and C3 represent the presupposition and the assertion of the cleft, respectively:

- (5-33) C1: $[e_1, j, a]$ cooked(e_1, j, a) \wedge cake(a)
 C2: $[e_2, x, b]$ cooked(e_2, x, b) \wedge cake(b)
 C3: $[s, m]$ be(s, x, m)

In the discussion that appears later, we will subsume both subtypes of relation under the umbrella category OLD.

Inferable Presuppositions

This category of presupposition is designed to capture cases where the presupposition is indirectly related to one or more propositions in the context, but has not a directly matching antecedent in the context at the time of interpretation. In such cases, the presupposition involved would be accommodated, in the manner described in chapter 4. The inferable category is intended to represent one of two ways in which such accommodation might take place, the second being described in the discussion of NEW presuppositions in the next section.

There are two ways in which antecedents can be created for presuppositions by means of an Inferable relation. The first way is by means of a LOGICAL RELATION between the presupposed proposition and a context proposition. The second is by means of the application of some fact or general knowledge schema that licenses an inference to the required proposition. We can capture this second type of Inferable relation in terms of

a set of axioms that are available to the interpreter of the cleft. The axioms that are relevant in describing inferential relations fall into two groups, as follows:

World Knowledge axioms where each axiom is based on speakers' knowledge about the state of the world or knowledge about other speakers; and

Metalinguistic axioms where each axiom is based on speakers' knowledge of discourses and the situations in which they take place.

We will describe both kinds of inference, those based on logical relations and those based on axioms, in turn.

Inferences based on Logical Relations

A particularly common presuppositional relation is one in which an antecedent can be formed for the presupposition from what Wilson [1975] terms a **BACKGROUND ENTAILMENT** of a proposition currently present in context. Such entailments are more general (i.e. less informative) entailments arising from other propositions.⁶ For example, the proposition (5-34a) might have the the background entailments given as (5-34b) to (5-34e) (in each case, glosses appear beneath the representation):

- (5-34) a [p₁, j, b] runs(p₁, j, b) ∧ business(b)
 (*John runs a business*)
 b [p₁, x, b] runs(p₁, x, b) ∧ business(b)
 (*Someone runs a business*)
 c [p₁, j, e₁] does(p₁, j, e₁)
 (*John does something*)
 d [p₁, x, e₁] does(p₁, x, e₁)
 (*Someone does something*)
 e [e₂] occurs(e₂)
 (*Something occurs*)

We can term each of these more general background entailments a **SUBSUMPTION** of the proposition in context, in this case (5-34a). How it relates to clefts can best be seen by

⁶We do not want to assume that all possible background entailments are generated at the point of the initial interpretation of sentences. Instead, they are generated only when required.

means of an example, such as (5-35). In this example, we give an utterance S1 from which the conditions C1 result, by entailment (although it is only really relevant that C1 occurs in the context, not how it arrived there). C1 describes a process of *running* that holds between John and a business:

- (5-35) S1: John runs a business.
C1: $[p_1, j, b]$ runs(p_1, j, b) \wedge business(b)

Given this context, it is possible to use a cleft S2, shown in (5-36). The presupposition of this cleft can be given as C2:

- (5-36) S2: What he does is write adverts.
C2: $[p_2, j, e_2]$ does(p_2, j, e_2)

At the time of interpretation, there is not a proposition in the context that can act as a directly-matching antecedent for C2. However, by going back into the discourse context, a proposition may be located from which a proposition that can act as antecedent can be derived. In this case, C1 is located as a proposition from which the appropriate proposition can be inferred by subsumption, as described above. The resulting proposition, a background entailment of C1, will provide a direct match for the presupposition C2. In this case, the appropriate background entailment will be C1', below:

- (5-37) C1: $[p_1, j, b]$ runs(p_1, j, b) \wedge business(b)
C1': $[p_1, j, e_1]$ does(p_1, j, e_1)
C2: $[p_2, j, e_2]$ does(p_2, j, e_2)

We can say that the presupposition of the *wh*-cleft in (5-36) relates to the propositions in context by means of the inferrable relation of subsumption. As we will see later in this section, the subsumption relation is particularly relevant in the description of *wh*-clefts.

While *do* is a common subsuming predicate, it is only able to subsume previous propositions that describe processes, not those describing events or states. For example, in (5-38), a *wh*-cleft attempts to use *do* to subsume a description of a state, that of John liking Duvel beer:⁷

- (5-38) S1: John likes Duvel beer.
S2: #What he does is drink a lot of it.

⁷For the purposes of the examples to follow, we will give the surface forms of the sentences from which the propositions arise, rather than the representations of the propositions concerned. It should be remembered, however, that the subsumption relation referred to in each case relates to the presupposition conveyed by the cleft, and the proposition (in each case, an entailment) of the context sentence.

(5-39) shows a similar situation where a *wh*-cleft with *do* tries to subsume an event description, the event concerned being that of the door slamming:

- (5-39) S1: The door slammed.
S2: #What it did was shut with a bang.

If *do* cannot be used for events, another subsuming predicate, *happen*, can. The use of this predicate is possible because the less informative entailment of event descriptions upon which it relies can be given as (5-40):

- (5-40) [e, x] happen(e, x)

Happen can therefore be used in context S1 given in (5-39), as (5-41) demonstrates:

- (5-41) S1: The door slammed.
S2: What happened was it shut with a bang.

We pointed out in chapter 3 that the *do*-insertion was not syntactically acceptable when the verb in the head of the cleft was a stative. For example, (5-42) was unacceptable:

- (5-42) *What Mary does is love ice cream.

As we noted in chapter 3, neither can *happen* appear with a sentential head that is the complement of a stative verb:

- (5-43) *What happens is that Mary loves ice cream.

This places obvious restrictions on the subsumption relation as far as statives are concerned: if the target proposition entailed by the context is a state description, it cannot be subsumed. We saw that this was the case for *do* in (5-40); (5-44) shows a similar lack of acceptability when *happen* is used for stative subsumption:

- (5-44) s1: John likes Duvel beer.
s2: #What happens is that he drinks a lot of it.

As far as the appearance of cleft presuppositions falling into this INFERRABLE category has been noted in previous research on clefts in discourse, Prince [1978:889] and Geluykens [1984:48] observe the appearance of what they term EXPLICIT information in *wh*-cleft *wh*-clauses, information which we have characterised here as being ENTAILED by

the discourse context. In addition, both remark upon the appearance of *happen* in *wh*-clefts, and Geluykens observes *do* to appear in these contexts, too. However, rather than providing a general explanation on the basis of entailment relations, as we do above, both Prince and Geluykens resort to independent pragmatic principles to explain these relationships. With respect to *happen* in *wh*-clefts, for example, Prince concludes

The antecedents 'something happened/seems to be happening' are not retrievable from anything in the preceding contexts; yet the discourses are coherent. There seems to be a pragmatic principle that says that events keep occurring—and that in our culture, at least, they are our proper and constant concern.

[Prince 1978:893]

Geluykens suggests a similar explanation for *do*:

It seems reasonable to assume that, just as the knowledge that events keep occurring is a pragmatic principle, it is also a pragmatic principle that 'people do things'. The fact that people perform all kinds of actions is probably always in our consciousness.

[Geluykens 1984:52]

On the approach advocated here, independent principles do not need to be invoked to explain the relationships described.

Inferences based on World Knowledge Axioms

Turning to world knowledge relations first, we can demonstrate the use of an axiom to license the accommodation of a presupposition using a simplified version of one of the corpus examples. In (5-45), S1 represents the current context, S2 a *wh*-cleft whose presupposition is inferentially related to that context:

- (5-45) S1: John has a gadget.
S2: What it does is remove fluff.

The content of S1 can be represented as the condition C1, as follows:

- (5-46) C1: $[s, j, a] \text{ owns}(s, j, a) \wedge \text{gadget}(a)$

The content of the cleft can be represented as C2 and C3 in (5-47), where C2 represents a presupposition that can be glossed as *it has some function*, C3 representing the assertion which we can gloss *that function is removing fluff*. In C3, the eventuality index e represents the process of fluff-removing, following the practice we introduced in chapter 4, of allowing predicative VPs to introduce eventuality indices. In a full representation, e would be decomposed further, but the partial representation below makes the point more clearly:

- (5-47) C2: $[s, x]$ has-function(s, a, x)
 C3: $[e]$ be(s, x, e)

On the approach we developed in chapter 4, the cleft presupposition in C2 has to be accommodated in the context C1. In the DRT model, we represented this by simply asserting the proposition into context, making it a *precondition* on the DRS. Here, however, we want to go more deeply into how such assertion can legitimately take place. In this case, in order to license the accommodation of C2, we can postulate a world-knowledge axiom known to the cleft interpreter which contains information to the effect that things have functions, unless it is specified otherwise. Such an axiom might be written as (5-48):

- (5-48) Function: $\forall y$ gadget(y) \supset has-function(s, y, e)

This axiom can be used to create an antecedent for the presupposition C2 if it can be instantiated to the values present in C2. In this case, y in (5-48) will be instantiated to a of the cleft presupposition C2, and e will be left uninstantiated, matching the uninstantiated x in C2. Because this match is successful—we can say it is possible to unify the two propositions, as observed above—we can license the accommodation of C2 as a precondition on a DRS; that is, we can assert it into context C1. The location and instantiation of axioms in this way can be seen as one way of licensing the accommodation of presuppositions such as C2, the instantiation process being the part of the procedure which we would describe as *INFERENCE*.

While we have postulated a quite specific axiom related to the function of *gadgets*, we can also suggest general axioms which reflect presuppositional relations that occur repeatedly in the corpus of data. Below, we give an informal rendering of each of these axioms, and give an example that belongs to the class covered by each.

Nine general axioms appear to cover the examples in sub-corpus B:

Means: For any state of affairs, there is a means of arriving at that state.

Result: For any process, there is a result.

Origin: For any entity or set of entities, there is an origin.

Source: If a speaker knows a fact, that fact has a source.

Need: For any two state of affairs, there exists something that is needed if it is to be possible to move from the first state to the second.

Cause: For any state of affairs, there is a cause.

Reason: For any state of affairs that has come about through deliberate human agency, there is a reason.

Affect: For any state of affairs, or any entity, about which they speak or in which they are involved, speakers and protagonists have feelings and opinions about that state of affairs or entity.

Opposition For any eventuality that holds or occurs, there exists a state of affairs in which an eventuality of the same type does not hold or occur.

Note that for an axiom to be selected, the state of affairs, entity, or eventuality quantified over in the initial clause of the axiom has to be present in the discourse context.

Examples exploiting each of the axioms in turn are as follows. (5-49) demonstrates a cleft S2 whose presupposition relates to the preceding context S1 by the relationship of MEANS:

- (5-49) S1: And here we have the tree.
S2: How I got to the tree was simply by adding one on each time.

(5-50) shows the relation of RESULT:

- (5-50) S1: If you just want to say look, these two descriptions simply aren't related to one another
S2: what you would get would be this (*points to diagram*)

The relation of ORIGIN is shown by (5-51):

(5-51) S1: Every transformation word I've heard is in at the moment is in the course for mathematics.

S2: That's where it all comes from.

(5-52) demonstrates the relation of SOURCE:

(5-52) S1: (*speakers are discussing the marriage of an acquaintance*)

S2: It was Serena who discovered she'd got married.

(5-53) shows a cleft whose presupposition is inferrable by virtue of the axiom of NEED. The context of this example is a meeting in which the speakers are discussing ways of improving some programming documentation:

(5-53) S1: I mean there's already some of that in the low-level programming documentation.

S2: What we actually need is a specification of what these things do.

The next relationship in the group is one of CAUSE, demonstrated in (5-54). In this example, the speaker refers to a state of affairs in which the expression *Caucasian* has come to mean *white-skinned*. The speaker goes on to name the cause of this state of affairs as *the Americans*, using the following cleft:

(5-54) It was the Americans who started that craze, I think.

In the following example, the speaker uses a cleft to give the REASON for her lateness:

(5-55) S1: I forgot then I remembered

S2: so that's why I'm about five minutes late you see

The relationship of AFFECT is demonstrated by the reverse *wh*-cleft in (5-56):

(5-56) S1: And these are people who have got quite a lot of money

S2: That's what annoys me

In this case, the cleft presupposition expresses the fact that the speaker has some feeling, namely annoyance. Both Prince and Geluykens observe the presence of relations like this category of AFFECT relations in *wh*-clefts. Prince observes:

In ordinary, non-ritualized discourse, it seems that the speaker's relevant thoughts, observations, opinions, reactions etc. (often negative) are taken to be the constant appropriate concern of the hearer.

[Prince 1978:891]

However, for none of the cases they observe do Prince or Geluykens suggest that the content of the cleft presupposition represents any particular relationship with a formal notion of context. Neither do they suggest that the relationships are relevant for any cleft type other than *wh*-clefts.

The relationship of OPPOSITION can be characterised as the principle that, for any state of affairs in which an eventuality (state, event, or process) can be said to hold or take place, there exist other states of affairs in which an eventuality of the same type does not hold or take place. For example, in (5-57), the context features a state of affairs in which an ordering event e_1 (in a restaurant) has taken place. The cleft S2 coheres by virtue of the absence of an event of the same type, an ordering event e_2 (note that the speaker is not saying that the *identical* event has not taken place):

- (5-57) S1: I've ordered.
S2: It's me who hasn't ordered.

(5-58) is a similar example featuring a *wh*-cleft:

- (5-58) S1: You could get to be a general by being an aristocrat
S2: but what you could not get to be was to be the general's chief
of staff.

As (5-57) and (5-58) show, this relationship may be accompanied by the appearance of negation in the cleft complement. The opposition need not, however, be expressed by explicit negation: it can also be effected by the presence in the cleft of a verb that is *lexically opposite* to one in the previous discourse. This variation in linguistic form does not affect the nature of the coherence relation. An example using lexical opposition rather than negation appears as follows:

- (5-59) S1: I quite like documentaries in general.
S2: What I loathe are those interminable heart transplants.

Inferences based on Metalinguistic Axioms

The second group of principles upon which relations between cleft presuppositions and discourse context are based are drawn from speakers' knowledge of the discourse and its situation. Following Prince [1978], we will term these METALINGUISTIC relations. It

is possible to distinguish four common means by which metalinguistic presuppositions are related to context, as follows:

Topic: If a speaker x_1 contributes to the discourse an utterance S1, S1 is about some topic x_2 .

Importance: If a speaker x_1 participates in a discourse, there exists some topic x_2 that is important to the speaker.

Speaker Meaning: If a speaker x_1 uses a linguistic form S1, x_1 means something by S1 that can be given as an alternative linguistic form S2.

Significance: A linguistic or semantic element x_1 has a significance x_2 .

The first relation is based on the principle that speakers have an idea of the TOPIC of what they are talking about—we can say that all discourse is *about* something. Presuppositions such as that appearing in (5-60) seem to be based on this principle:

(5-60) It's people who are still working on their PhDs that we're talking about here.

A second, related type of metalinguistic relation is based on the perception that speakers consider particular topics important, and some more so than others. Clefts whose presuppositions rely on this principle of IMPORTANCE appear in (5-61) and (5-62):

(5-61) S1: I don't like this business of dear old England is full of eccentrics
—let's encourage people to be funny individuals and things like that
S2: I mean what really matters is systematic collective activity

(5-62) S1: It isn't a business of learning—well yes there are techniques to learn but
S2: it's very much the application or the criticism of the very techniques themselves which is the important thing

The third and fourth types of relationship that belong in this metalinguistic category both concern kinds of *meaning* relation. The third relation, which we will term SPEAKER MEANING, concerns the assumption that when a speaker says something, he or she means something by saying it. The axiom of speaker meaning describes the fact that a speaker can always clarify what he or she means by some previous utterance by means of an alternative linguistic formulation. An example based on this principle appears in (5-63), where S1 and S2 are context sentences, S3 the cleft:

- (5-63) S1: you turn left and immediately right again in order to go through Farnham
 S2: Farnham actually you go around—it's got a bypass—you don't go into Farnham
 S3: No what I mean is you cross a major road

The fourth and final metalinguistic relationship discovered also tends to be realised by the verb *mean*. This relationship, which we will term SIGNIFICANCE, can be used both literally and metaphorically. In the literal case, it relates a linguistic item to its meaning. This is distinct from the case described above, where meaning was attributed not to linguistic items alone, but to speakers. An example of the literal use of the significance relation appears in (5-64):

- (5-64) S1: I want to be able to talk about elements being prime.
 S2: And what prime means is that anything below this description has that primitive in it.

The metaphorical use of the significance relation occurs when some element—which may be an eventuality or an entity—is said to have some other significance of meaning in the symbolic sense. An example appears in (5-65):

- (5-65) Of course what this means is that Sellers is for the high-jump

We can summarise the Inferrable relations suggested here as follows:

- Logical relations
 - Subsumption
- World-Knowledge relations
 - Function
 - Means
 - Result
 - Origin
 - Source
 - Need
 - Cause

- Reason
- Affect
- Opposition
- Metalinguistic relations
 - Topic
 - Importance
 - Speaker Meaning
 - Significance

While we do not want to assert that the list above is exhaustive, it serves to capture the Inferrable presuppositional relations encountered in sub-corpus B.

New Presuppositions

In cases where presuppositions are NEW, the eventuality described by the cleft presupposition has no propositional antecedent currently in the discourse context, and neither can one be inferred for it on the basis of any inferential relation such as those described above. In these cases, we describe the presupposition as being NEW in the context.

Of course, it is difficult to give examples of this type, since whatever the quantity of context that is included it is possible to suggest that some coherent relation occurs on the basis of context prior to that given. While we do not have the space here to provide the full context of (5-66), for example, we can explain that nowhere in the previous context does a proposition arise that can be glossed as *someone is the editor of the Tundraland and Westingham*. Although we could postulate the existence of an axiom such as *for any newspaper x , has-editor(x)*, newspapers are not present in the discourse context—and so, on the stipulation given above that the referent of the antecedent clause of the axiom should be present if that axiom was to be selected as the basis of an inferential relation, the presupposition cannot be classed as Inferrable, either. The cleft presupposition in (5-66), therefore, seems to entirely lack a relationship with the preceding context:

- (5-66) S1: Simon Crawley ...he's the Gallic person yes yes
 S2: Now is it Simon Crawley who's the editor of the Tundraland
 and Westingham

A similar example appears in (5-67), simplified from one that appears in the corpus for the sake of the analysis:

- (5-67) S1: An unpleasant situation has arisen.
 S2: It's this that makes an absolute fool of Mallet.

Again, although some of the individual entities referred to in S2 may have been already mentioned in the discourse (it is likely, for example, that *Mallet* has already been referred to) the coherent relation between the cleft presupposition, which can be glossed *something makes a fool of Mallet*, and the previous context, is entirely absent. That is, the presupposition of S2, represented below as C2, is has no propositional antecedent in context C1, and no relevant or plausible axiom exists that could be instantiated to provide one:

- (5-68) S1: An unpleasant situation has arisen.
 C1: $[s_1, s_2]$ exists(s_1, s_2) \wedge situation(s_2) \wedge unpleasant (s_2)
 S2: It is this that makes a fool of Mallet.
 C2: $[s_3, m, x]$ makes-a-fool-of(s_3, m, x)

The cleft, however, is acceptable. In these cases, we would argue that the content of the cleft presupposition is accommodated simply by asserting that content into the discourse context.

Having introduced the statuses of presupposition that can occur, we can turn to the range of these presuppositions that appeared in the data.

5.3.3 The Distribution of Presuppositional Relations

The basis for the study of presuppositional relations was the 150-example sub-corpus B. The frequencies of presuppositional relations across cleft types, significant to 0.05 according to the chi-square test, were as shown in figure 5.5. In the analysis to follow, we will be particularly concerned with the distinctions between clefts in terms of which cleft type is favoured given a particular presuppositional relation. Of the tendencies that appear in the data, there are three that are particularly striking:

Presuppositional Relation		<i>it</i> -clefts	<i>wh</i> -clefts	reverse <i>wh</i> -clefts
Old		16	3	2
Inferable	Subsumption	0	15	5
	World Knowledge	13	17	8
	Metalinguistic	1	7	1
	Affective	4	8	6
New		16	0	28

Figure 5.5: Frequencies of Presuppositional Relations by Cleft Type

1. *Wh*-clefts do not carry presuppositions that are New in relation to context.
2. *Wh*-clefts are particularly favoured for relations of Subsumption, and for Metalinguistic relations.
3. *It*-clefts are particularly favoured for Identity relations.

We will address each cleft type in turn.

The Presuppositions of *Wh*-clefts

Sornicola [1988:366] observes two distinct presuppositional relations for *wh*-clefts, one in which an antecedent for the presupposed material appears in the context, one in which it does not. In fact, slightly more can be said about *wh*-clefts than this: we can show that the absent antecedent is in fact always inferable from the context. The first point to note in support of this is that there appears to be a real constraint on the appearance of propositions unrelated to discourse context as *wh*-cleft presuppositions. We can show that this is not just a quirk of the data by comparing examples of *wh*-clefts with *it*-clefts, which can carry such presuppositions. An example appears in (5-69). Both the *it*-cleft and *wh*-cleft convey presuppositions unrelated to context either by entailment or inferential relation:

- (5-69) A: Sue, this is Marta.
a B: Hi! It was your flat I rented when you were in the States.
b B: # Hi! What I rented when you were in the States was your flat.

Additional support for this view comes from data presented by Prince [1978:895]. She draws attention to examples in which the cleft head is part of an idiom. She points out that where *it*-clefts are acceptable in this role, *wh*-clefts are not. Her examples are as follows:

- (5-70) a It's sort of an arbitrary line that you're drawing.
 b #What you're drawing is sort of an arbitrary line.
- (5-71) It's obvious that I'm a woman and that I enjoy being a woman. I'm not overly provocative, either.
 a It's the thin, good-nigger line that I have to toe.
 b #What I have to toe is the thin, good-nigger line.

In each case, the *wh*-cleft conveys a clear impression that the hearer is expected to be acquainted with the information in the presupposition prior to the time of utterance of the cleft. We would suggest that this demonstrates a constraint on accommodation: an unrelated presupposition cannot be accommodated if it is presented in a *wh*-cleft.⁸ When it comes to why such a constraint should exist, we can only propose that the combination of leftmost position in the cleft construction, traditionally the site of information already available to the hearer or reader (cf. Halliday [1967]), and the fact that this position in the *wh*-cleft is marked as the site of presupposed information, together represent too strong a signal that the information is in some sense contextually available for New information to be sited there.

A second point to note with respect to *wh*-clefts is that they are the most frequent choice for subsumption relations, such as those appearing in the clefts in (5-72):

- (5-72) S1: John's had an accident.
 S2: What happened was he fell off his skateboard.
- (5-73) S1: John's made a grand gesture.
 S2: What he's done is join the Foreign Legion.

Note that in the above cases the cleft head is either a sentential constituent or a non-finite VP. From this syntactic fact arises the first constraint on the use of some cleft types for this relation, since, as we noted in chapter 3, sentential constituents as subjects

⁸Prince [1978:892] observes that *wh*-cleft heads must not present information that is 'more prominent in the hearer's consciousness than the information in the *wh*-clause'. This is a relative constraint between the roles of head and complement which we discuss in the next section.

are not acceptable in *it*-clefts and reverse *wh*-clefts. This prevents subsumption relations using *happen*:

- (5-74) a *It was that John fell off his skateboard that happened.
b *That John fell off his skateboard was what happened.

While *wh*-clefts, and to a lesser extent *it*-clefts, can specify similar content to that specified by a sentential constituent using *that* or *this*, this is obviously only appropriate in discourse situations where anaphora is possible. This is likely to be only a few cases, since the main function of clefts based on subsumption relations is to convey a further explanation of a known eventuality, the explanation commonly being new information. The notion of new information as it relates to discourse objects other than eventuality descriptions can be intuitively appreciated for the purposes of the current discussion as 'not currently available in the discourse context'. We define the notion more formally for the purposes of the discussion of information structure in section 5.4 below.

A second factor involved in the preference for *wh*-clefts when presuppositions are required to subsume existing propositions in context concerns the opportunities only the *wh*-cleft offers for the serial ordering of the information involved. There are two strands to the serial ordering argument, one of which has been discussed in section 5.2, the other of which appears in more detail in section 5.5.

First of all, in cases where the presupposed information is a subsumption, the remainder of the information is likely to be new, consisting of a further explanation or elaboration of the subsumed eventuality. This in turn suggests that the New information will be quite lengthy. As Geluykens [1984] has pointed out, and as we will see below in the discussion of information structure, both *it*-clefts and reverse *wh*-clefts tend to have short head constituents. Geluykens suggests that this is due to speakers' reluctance to postpone the 'Given' information that appears in the cleft complement. While we do not adopt the assumption that Given information necessarily appears in cleft complements, it does seem that there is evidence to support the view that cleft *presuppositions* are processed first by humans, the corollary of this being that a processing load results from postponing the presupposed information. We will give the full argument behind this view in section 5.5; for the moment, it is sufficient to note that the *wh*-cleft, being the sole cleft type that is able to place information serially *after* the presupposed informa-

tion, will be the best candidate for supplying large quantities of New, non-presupposed information while minimising processing load.

The second strand to the serial ordering argument is one that has already been discussed in section 5.2. This is the principle that, if antecedent-requiring material, such as a subsuming presupposition, exists in one part of the cleft and New information in another, then it maximises ease of processing to place the lexical material that requires an antecedent nearest to that antecedent, and the New information that is presumably to be enlarged upon later nearer to the subsequent lexical material that constitutes that thematic development. Again, the *wh*-cleft is ideal for this task in the case of presuppositional relations of subsumption: presupposed information is nearer to its antecedent, and the part of the cleft carrying New information is nearer to the part of the discourse that refers to it subsequently.

The final regularity to discuss is a related one, to the effect that *wh*-clefts are also preferred when the presupposed information is metalinguistically related to discourse context. Often, clefts with metalinguistic presuppositions are used to introduce a New topic, or to structure the discourse in a way that requires a bulk of New information to be provided. An example appears in (5-75):

- (5-75) What I want to talk about is basically a way of creating big *Is* out of smaller ones.

Compare the choice of *wh*-cleft to the corresponding *it*-cleft and reverse *wh*-cleft:

- (5-76) a ?It's a way of creating big *Is* out of smaller ones that I want to talk about.
b ?A way of creating big *Is* out of smaller ones is what I want to talk about.

The *it*-cleft and reverse *wh*-cleft are more clumsy in general on the basis of the ordering principles discussed above. Again, as we saw in section 5.2, the ordering of the message needs to take into account what part of the message is going to be related thematically to preceding discourse, and what to succeeding discourse, and arrange the information accordingly.

The Presuppositions of *It*-clefts

It-clefts seem to occupy almost the complete range of presuppositional relations. One exception in the data is the non-appearance of *it*-cleft presuppositions which are subsumptions of previous context, which were discussed above. Unlike the case of *wh*-clefts, the range of syntactic head types required—sentential constituents and non-finite VPs—are often marginal in *it*-clefts. In addition, the amount of New information that generally has to be conveyed by a cleft with a subsuming presupposition causes problems for the *it*-cleft in relation to the principles of serial ordering discussed above. *It*-clefts simply present information in the wrong order to be of use for the functions performed using a subsuming presupposition.

New presuppositions, on the other hand, are reasonably popular in *it*-clefts, but compared with the other cleft types, the preference for identity relations is the most significant. We can explain this tendency in terms of the principle of EFFICIENCY suggested by van der Sandt [1988], one of the set of principles we discussed in chapter 4, which together constrain the incorporation of propositions into context.

The Principle of Efficiency is intended to prevent the use of repetition and elaborate phrasing that may give rise to implicatures that less complex structures would not. In the case of the identity relation, the presupposition is composed of information already entailed by the context. This could lead to repetitiveness in the cleft complement, as in the following examples of each type of cleft:

- (5-77) A: What caused the earthquake?
a B: What caused the earthquake was the San Andreas fault.
b B: The San Andreas fault was what caused the earthquake.
c B: It was the San Andreas fault that caused the earthquake.

The *it*-cleft is unique in that it allows omission of the complement clause in cases of redundant repetition, as follows:

- (5-78) A: What caused the earthquake?
B: It was the San Andreas fault ϕ .

The possibility for complement omission would contribute to the preference for *it*-clefts when the presuppositional relation is an identity relation.

Complement omission does not take place, however, when the propositional antecedent in the context is not currently active in the discourse context, for reasons of discourse structure or the passage of time. In addition, *wh*-clefts and reverse *wh*-clefts are sometimes used when the implicatures created by repetition are required. For example, the repetition of the part of the *wh*-cleft that conveys an entailed presupposition can create a pedantic or emphatic effect. Such repetition conveys a speaker's refusal to adopt material suggested by another speaker, and hence promotes a feeling of social distance or even disagreement:

- (5-79) A: What caused the problem?
a B: It was a glitch in the power supply ϕ .
b B: What caused the problem, Jones, was your own stupidity.

Most notable about *it*-clefts, however, is the fact that the data split roughly equally across all three types of presupposition: New, Inferrable, and Old. This marks the *it*-cleft as the most flexible of all three cleft types in relation to the presuppositional relations that appear.

The Presuppositions of Reverse *Wh*-clefts

Reverse *wh*-clefts, like *it*-clefts, have a broad range of presuppositional relations open to them, but differ from *it*-clefts in that over half the presuppositions fall into a single category, namely New. This appears to be due to the fact that, as we shall see in the discussion of information structure below, reverse *wh*-clefts can only have old information appearing in head position (again, this can be intuitively recognised as 'information currently present in the discourse context'; we define it formally in section 5.4). This means that any novel contribution is likely to appear in the presupposition. This same fact may also account for the low frequency of Old presuppositions: if heads are old, a cleft with an Old presupposition as well will have a very low level of informativeness. Although clefts with this information structure do occur, they are appropriate in only a limited range of situations.

Presuppositional Relation		<i>it</i> -cleft	<i>wh</i> -cleft	reverse <i>wh</i> -cleft
Old		yes	yes	possible but rare
Inferable	subsumption	no	yes	possible but rare
	world knowledge	yes	yes	yes
	metalinguistic	yes	yes	yes
New		yes	no	yes

Figure 5.6: Range of presuppositional relations by cleft type

5.3.4 Clefts and Presupposition: Summary

It was claimed at the beginning of the section that cleft presuppositions inhabit a continuum of relationships with previous context, and occupy overlapping but distinct ranges along that continuum. With respect to discovering where on that continuum the various cleft types lie, the examination of the data above revealed two apparently categorical constraints, namely that *wh*-cleft presuppositions could not carry descriptions of New eventualities, and *it*-cleft presuppositions could not act as subsumptions. Other tendencies were due to the differing possibilities the clefts offer for complement omission to avoid repetition, principles governing the serial ordering of information, and constraints on the usefulness of clefts combining a particular presuppositional relation with a particular discourse status of cleft head. This combination of tendencies and constraints is shown in figure 5.6.

Having examined the relationships that hold specifically between cleft presuppositions and the discourse context, we can turn to the related issue of the information structure of clefts as a whole, incorporating both the presuppositional relations described in this section, and the information asserted by the cleft.

5.4 The Information Structure of Clefts

In chapter 2, we examined claims that have been made in the literature with respect to the INFORMATION STRUCTURE of clefts, typically interpreted as their arrangement of Given and New information, following Halliday [1967]. In that chapter, three hypotheses related to information structure were reviewed in particular:

- Cleft sentences provide a means of arranging Given information before New information.
- Cleft sentences provide a means of placing New information towards the end of the sentence.
- Cleft sentences provide a syntactic means of signalling a division between the sites of Given and New information in the sentence.

In some cases, a stronger version of the third hypothesis was advanced, as follows:

- Cleft sentences provide a syntactic means of signalling a division between the sites of Given and New information in the sentence, specifically by placing New information in head position, and Given in complement position.

Although examples were available to support all three hypotheses, it was possible to show that counterexamples existed for each one. In particular, we saw examples of cases in which:

- New information appeared before Given in a cleft sentence;
- New information was not at the end of the sentence;
- a combination of Given and New information appeared in a single syntactic part of the cleft—i.e. head or complement; and
- New information appeared in the complement, and Given in the head.

The effect of the discussion in chapter 2 was therefore to reveal a certain amount of chaos with respect to the information structure of clefts. In particular, two points were demonstrated: one, that no single claim regarding information structure could generalise across clefts of all three syntactic types, and two, that examples of the same cleft type could have different information structures.

The purpose of this section is to show that, in spite of the perhaps unanticipated variety that exists with respect to the information structure of clefts, some important regularities do exist, although these are not necessarily of the same nature as those suggested previously. In particular, we will attempt to answer the following three questions:

- Are there any generalisations about information structure that apply across all three types of cleft?
- Are there any characteristic information structures that exist within cleft types?
- Related to the previous question, what are the distinctions across cleft types with respect to these characteristic information structures?

In what follows, it will be suggested that generalisations of all three kinds can usefully be made. First of all, however, we need to establish a simple formal framework for talking about information structure, in order to make perfectly clear the meaning of the terms appealed to throughout this section, and the status of the conclusions reached.

5.4.1 Analysing Information Structure

Although other taxonomies do exist of the kinds of discourse status entities can have (cf. Prince [1981], for example), the previous research on the information structure of clefts discussed in chapter two adopts almost uniformly the GIVEN-NEW dichotomy, originating with Halliday [1967].

Halliday's Given/New Dichotomy

Halliday's description of Given and New information [1985:27] can be paraphrased as follows:

Given information Information that is presented by the speaker as recoverable to the listener. It may have been mentioned before, it may be something that is in the situation, or that is not around at all but the speaker wishes to present it as Given. The meaning is: this is not news.

New Information Information that is presented by the speaker as not recoverable to the listener. It may be something that has not been mentioned, it may be something unexpected, whether previously mentioned or not. One form of Newness is contrastive emphasis. The meaning is: attend to this, this is news.

The basis of the relationship between Given and New and the sentences they are intended to describe is the TONE UNIT, an abstract phonological unit⁹ usually characterised as having one particularly prominent accent—the NUCLEUS—and a recognisable intonation contour. On Halliday's account, tone units function as the realisation of an abstract unit of INFORMATION, which he terms the INFORMATION UNIT or IU (Halliday [1976:202, 1985:275]). It is in relation to these units, rather than in relation to the tone unit, that information structure is assessed.

According to Halliday, IUs may consist of either a Given and a New element, or be composed entirely of New information. New information is marked by the nucleus, and in general, Given information precedes New.¹⁰ Typically, the position of the nucleus marks the end of the New information, which is most usually the element at the end of the IU. However, in cases where the nucleus does not occur on the final element of the IU¹¹, any material to the right of the nucleus is taken to be Given, whether that material is accented or not.

Influential though Halliday's account of information structure has been, it will not be adopted here for the reason that several sources of indeterminacy exist within it, some of which Halliday himself points out. First of all, Halliday [1985:276] notes an inherent difficulty in the Given/New framework in terms of deciding on the *extent* of the New information. Although the accented element can usually be understood to be the culmination of New information, where that information begins is more problematic. Halliday [1985:276] indicates this indeterminacy through the following two examples. In (5-80), the context suggests that New information extends from the accented item *love* throughout most of the clause, with the exception of *it*, which is Given (since it repeats material from the immediately preceding context):

(5-80) I'll tell you about silver. It needs to have LOVE.

⁹Crystal's notion of tone group discussed above in section 5.1 is similar, if not identical, to Halliday's notion.

¹⁰The element carrying the nucleus is said by Halliday to be carrying information focus, which is claimed by Halliday to mark the most informative part of the information unit. We will not be making use of this notion of focus, since it is the full extent of the New information, rather than the position of its most informative peak, that is of interest here.

¹¹Halliday's terms the appearance of the nucleus on elements other than that appearing finally in the IU 'marked information focus'.

In (5-81), the identical accentual structure is placed in a context that indicates a Given analysis for the entire clause except for the accented item:

(5-81) I'll tell you what silver needs to have. It needs to have LOVE.

Although Halliday does not use the term himself, this issue of how the SCOPE of accents can be delimited, addressed at length by Selkirk [1980], Gussenhoven [1983] and Taglicht [1984] among others, is still an open research problem in phonological theory.

A second difficulty comes back to Halliday's definitions of Given and New information, paraphrased above. In both cases, the definition relies to some extent on what speakers *choose to present* as Given or New, and not on the actual status in the discourse context of any element of which Given or New status is predicated. The appearance of such imponderables as what speakers are intending renders extremely difficult the analysis of utterances into Given or New parts.

Thirdly, Halliday's Given/New distinction is inextricably involved with the accentual patterns of utterances. Problems arise for the analyst both from this fact, and from the way Halliday's categories relate to accentual pattern. On the first point, it is clear that, if written language is to be analysed, some metric other than accentual pattern needs to be found. As it happens, the data examined in this chapter is exclusively spoken data, but it would be preferable if the account developed for it could be generalised to apply to written data and to casually-recorded spoken data for which the phonological diacritics are perhaps less than reliable. As for the second point, the analysis Halliday suggests runs into technical difficulties in that it does not consider sentence accents other than nuclei to be indicators of New information. As we saw in chapter 4 in our examination of Chomsky's notion of Focus, the role of subsidiary accents seems at least in some cases to be comparable to that of nuclei in indicating the status of information.

The fourth and final problem to note is that it is not clear what kinds of elements the status Given and New applies to. In the analyses he gives, Halliday appears to apply the status Given and New to the surface structure of discourse, indicating information status boundaries to be co-extensive with syntactic boundaries. As Gussenhoven [1983] has suggested, it appears to be more appropriate to apply such an analysis to the semantic objects that are realised by surface structure, for two reasons. The first is that the

evidence from indicators of information structure such as sentence accent suggests that different surface structures realising the same semantic content should be accorded the same, and not different, information status. The second is that, as was demonstrated in chapter 2, semantic features, such as negation or femaleness, can be legitimate components of the ‘New’ contribution of an utterance. On this basis, it would seem to be preferable to represent information structure at the level of a semantic representation of the discourse and context, rather than at the level of syntax or surface realisation.

The points above can be summarised simply as the objection that Halliday’s theory of Given/New, although intuitively appealing and undoubtedly influential, is difficult to apply to a range of everyday data in a replicable way. In this section, therefore, we will adopt an alternative analysis. The analysis suggested may be over-simple, but it suits immediate purposes and has the advantage of being applicable to the formal model of context developed in chapter 4.

A Framework for Analysing Information Structure

In this section, we look at how discourse status can be assessed with respect to the objects that populate our DRT model of context, namely discourse MARKERS. Some markers stand for entities, others for eventualities, as we saw in chapter 4. The treatments we propose for each type of object are distinct, but related.

Entity Markers

The treatment of entity markers adopted here consists simply of two discourse statuses, OLD and NEW, decided on the basis of the linear DRT representation of the content of the sentences involved. Which elements have which status is decided on the basis of a comparison between the incoming proposition and the previous context. To take a simple example, we can assess the sum of New information contributed to context C1, which we could gloss as *someone walked*, by sentence S2, as follows:

- (5-82) C1: [e, x] walked(e, x)
 S2: John walked.
 C2: [e, j] walked(e, j)

On the basis of the comparison of C2 with C1 we can say that the New element is the one that does not appear in prior context, namely *j* (or John).

At this point, a brief observation is in order relating to the notion of CONTRAST. We do not in this thesis go into this issue in much detail, so we will simply assume a tenable position on the subject. Following Bolinger [1961], Ladd [1983] and others, we take the view that heightened contrastive effect is a result of the narrowness of the scope of New information in the sentence in question. The more New information that is conveyed by the sentence, therefore, the less the contrastive effect. We can illustrate this by a couple of examples. In (5-83), the context sentence S1 appears to require a response, given as S2, which is composed entirely of New information. The result is a non-contrastive reading of S2:

- (5-83) S1: What happened?
S2: A car slid off the road.

In (5-84), on the other hand, the scope of the New information in the sentence S2 can be described as narrow, since only the object referred to by the expression *a car* is New:

- (5-84) S1: Did a bus slide off the road?
S2: No, a car slid off the road.

The result in (5-84) is a contrastive reading of S2. These observations are relevant to clefts in that constituents in which a small amount of New information appears to be confined, such as the heads of *it*-clefts and reverse *wh*-clefts, tend always to get a contrastive reading. The heads of *wh*-clefts, however, which are composed of longer and more complex descriptions, do not receive such readings so frequently (although they may, particularly if only part of the information conveyed by the cleft head is New). In both cases, however, note that only New information can receive a contrastive reading.

To return to the discussion of the discourse statuses of cleft heads, it remains to note that in some cases, a third discourse status seems to be required in addition to Old and New, described above. In these cases, the element specified by the cleft head is strictly New on the definition above, but is inferrably related to a marker that is currently present in the context. Examples are the relationship between the referents of *wall* and *house*, or *mountain* and *mountaineer*. This type of relationship has been captured by Prince [1981] in her INFERRABLE category, and noted by others, such as Hirst [1981:27],

in terms of sets of particular semantic relations where the elements involved are judged to be inferrably related to one another. It is still an open question what the precise relationships in such a taxonomy of inferrable relations should be, but Hirst suggests the following might be relevant:

- part of
- subset of
- aspect of
- attribute of
- closely associated with

A conclusive analysis of which elements are inferrable from the context and which are not is beyond the scope of this research, and so this third status will be mentioned only in cases where it appears particularly relevant to differentiate it from our New status. For the most part, however, in the analysis to follow, inferrable elements will be equated with New information, on the simple criterion that inferrable elements, like New elements, lack identical antecedents in context.

Eventuality Markers

One shortcoming of existing taxonomies of information status (cf. in particular Prince's otherwise excellent taxonomy [Prince 1981]) is that they are rarely if ever directed towards an analysis of discourse objects other than entities, such as eventualities. Our practice of introducing markers for such discourse objects means that they can be treated similarly to entity markers, using the set of discourse statuses developed in the previous section for talking about the status of presuppositions in the context. For example, if an eventuality index stands for an eventuality that is unrelated to previous context, we can treat this as a New eventuality in the manner described in section 5.3. The eventuality concerned may of course have Old participants, as in the following example:

- (5-85) S1: Mary walked.
C1: $[e_1, m]$ walked(e_1, m)
S2: Mary ran.
C1: $[e_2, m]$ ran(e_2, m)

In this case, we can say that an Old participant, m , is taking part in a New event, e_2 . If a marker stands for an eventuality that is related by axiom to previous context, we say it is inferrably related, in the manner described in section 5.3. Finally, markers standing for eventualities whose descriptions arise directly from explicit utterances, and are thus present in the context without inference, are Old, as before.

Since the participants involved in the eventuality may be of a different status to the eventuality itself, guidelines for compositionality—that is, the status an eventuality can be expected to have given the status of its participants—need to be set out for these cases. Whether two eventualities can be the same if they have different participants is an open philosophical question, and we do not expect to be able to solve it here. Rather, a reasonably coherent position needs to be adopted which we can use for the sake of the analysis.

In fact, it appears to be possible for two descriptions to be interpreted as alternative views of the same eventuality, even if the descriptions have no participants in common, even the identification of the action that has taken place. It is this that licenses disagreements such as the following:

- (5-86) A: And then John ate the last marzipan carrot.
B: No! Mary threw the pizza over Mike.

Here, we can say that every aspect of the eventuality is under dispute between A and B, except that fact that some eventuality actually took place, the eventuality being defined by a position in space and time. On this basis, there is no reason to suppose that a whole set of New participants in an eventuality description will necessary constitute a New eventuality.

However, in the discussion in section 5.3 we defined Old relations as cases in which the presupposed proposition was already ENTAILED by the context. In the case of B's utterance in (5-86), it is clear that her contribution is not an entailment of the context, even though they may be debating the 'same' eventuality. We will therefore take the position that, for an eventuality to be Old, the content of the eventuality description has to be entailed by the context. This in turn implies that all the participants in the eventuality will also be Old. This appears to be a coherent position to take for the few cases where the issue arises in the analysis.

This framework for discourse status just described enables us to capture some important regularities about the information structure of clefts. As mentioned in the introduction to this section, these fall into three particular types: regularities across clefts in general, characteristic information arrangements within cleft types, and related distinctions between cleft types. We will look at each of these kinds of regularity in turn.

5.4.2 General Features of Cleft Information Structure

The Correlation between Cleft Syntax and Information Structure

In chapter 2, we reviewed claims that the standard information structure for clefts of all kinds was either Given in the complement and New in the head, or at least some syntactic division between the two. As noted above, counterexamples were available for each of these claims, and we will briefly note here the prevalence of those counterexamples according to the data in sub-corpus B. Although Given and New are not the terms we use here, the framework for the analysis described above does turn up the required data if we concentrate on the appearance of New information.

The weak version of the claim that the cleft syntax serves to indicate information structure is that Given and New information are at least divided from one another. Counterexamples to this would be cases in which New information appeared in both the head and the cleft complement. The evidence against this claim is convincing: 14% of reverse *wh*-clefts, 60% of *it*-clefts, and 92% of *wh*-clefts examined contained at least one New element in both head and complement.

Among the counterexamples to the strong version of the claim, that New information appears in cleft heads and Given in cleft complements, are those clefts in which New information appears in complements. Again, the data are resoundingly against this conclusion: 90% of *it*-clefts, 92% of *wh*-clefts, and 100% of reverse *wh*-clefts contain at least one New element in the complement or *wh*-clause. Note that this underlines the fact that appearance of New information in cleft presuppositions is systematic and prevalent, rather than being the exception. We cannot therefore expect either claim to be of use in an account of clefts in discourse.

Two General Claims

There are two features observable with respect to the information structure of clefts which appear to be general across all three cleft types. It is possible to express these as well-formedness requirements on clefts, as follows:

- A minimum requirement for INFORMATIVENESS exists, namely that the cleft copula should introduce a New eventuality description.
- A minimum requirement for COHERENCE exists, namely that no cleft can carry entirely New information.

We will address each of these two requirements in turn.

The Informativeness Requirement

In chapter 4, we listed van der Sandt's four constraints on integrating propositions into discourse context. One of them, the principle of INFORMATIVENESS, was summarised as follows:

Informativeness: the information contributed by the sentence must not be a subset of that already in context.

The informativeness principle simply ensures that each proposition to be integrated into context has some New content. It implies, for example, that if the presupposition borne by a cleft contains no New information, it will not be re-integrated into current context. This prevents examples such as the following unacceptable cleft being generated or accepted:

(5-87) #John walked and it was John who walked.

In order for a cleft to be acceptable, however, it is not necessary that any of its individual participants are New to the context. For example, (5-88) is acceptable, even though the presupposed information *someone walked* is Old in the context, as is the entity *John*:

- (5-88) I knew someone walked, and I saw John a moment ago, but I didn't know it was John who walked.

The difference between the acceptable (5-88) and the unacceptable (5-87) is the MINIMAL INFORMATIVENESS REQUIREMENT for clefts mentioned above, namely that the *relationship between* the referent of the cleft head (in this case *John*) and the existentially-quantified variable in the cleft presupposition must not already be part of the context. We can capture this by stipulating that the state description introduced by the copula, which represents this relationship, must be a New element. This ensures that each cleft performs at least the function of conveying a novel conjunction of the element specified by the head and the proposition specified by the complement.

The unacceptability of example such as (5-87) suggests something important about the function of clefts, namely that they serve to convey exactly such novel conjunctions between head elements and the presuppositions whose variables the element specified by the head instantiates.

The Coherence Requirement

A second well-formedness constraint seems to exist, on clefts at least, to ensure that each cleft expresses some coherent link with the previous discourse in the form of Old information. It is possible to show as evidence for this that clefts composed entirely of New information are unacceptable, even in contexts where all-New information is acceptable in the case of declaratives. An example of such a context is the traditional question *what happened?*:

- (5-89) What happened?
a A cat's stuck up a tree.
b # It's a cat that's stuck up a tree.
c # A cat is what's stuck up a tree.
d #What's stuck up a tree is a cat.

To look more closely into the nature of the requirement for explicit coherence between clefts and previous context, we can examine the following example:

- (5-90) #Hi! It was in Dunfermline I got my gerbil.

In (5-90), both the eventuality specified by the presupposed information and the element specified by the cleft head are apparently New (even though the participant denoted by *I* in the presupposition is arguably Old, on the basis that the speaker is contextually available).

As we saw in section 5.3, there is no general constraint on presupposing information that is unrelated to the previous context. We can demonstrate this by constructing a context for (5-90) in which the head element, but not the presupposed information, is related to previous context. The result, shown in (5-91) below, is perfectly acceptable:

- (5-91) A: I've just come back from Dunfermline.
B: Really? It was in Dunfermline I got my gerbil.

(5-90) would also be acceptable in a context where the head element was New, and the eventuality described by the presupposition Old:

- (5-92) A: Where was it you went to get your gerbil in the end?
B: It was in Dunfermline [I got my gerbil].

What seems to be wrong with the example in (5-90) is not that its head specifies a New element, or that the eventuality described by the presupposition is New, but that both occur at the same time. The problem is the *combination* of New information in the complement and a New eventuality in the presupposition. We can see this is general across all three types of cleft. (5-93a) and (5-93b) are the closest *wh*-cleft and reverse *wh*-cleft equivalents of the *it*-cleft in (5-90):

- (5-93) a # Hi! Where I got my gerbil was Dunfermline.
b # Hi! Dunfermline was where I got my gerbil.

The minimal coherence requirement therefore seems to be that either the cleft head may specify a New element, or the cleft presupposition may specify a New eventuality, but that no cleft may do both at the same time.

On the basis of this, we can state a MINIMAL COHERENCE REQUIREMENT for clefts, as follows:

Coherence: No cleft can both specify a New head element, and presuppose a New eventuality.

5.4.3 Features of Information Structure Specific to Cleft Types

In this section, we will look in general at the patterns of Old and New information that occur in all three kinds of clefts, and state some constraints on information structure specific to individual cleft types. We will examine *it*-clefts, *wh*-clefts, and reverse *wh*-clefts respectively in the sections below.

The Information Structure of *It*-clefts

As Prince [1978] has pointed out, the information structure of *it*-clefts falls into two major patterns. Prince's terms STRESSED-FOCUS and INFORMATIVE PRESUPPOSITION *it*-cleft and their definitions are discussed in chapter 2, and we will not repeat that discussion here: suffice to say that the two information structures we note are similar to Prince's types, if defined more formally.

The first major *it*-cleft information structure is one in which the cleft head specifies a New element, and the eventuality specified by the presupposition is Old. In this case, the confinement of New information to the cleft head has the effect of producing contrast, which, as Werth [1984] points out, is itself a coherent relationship with preceding discourse. Clefts such as (5-94) fall into this category, where the head of the second speaker's cleft supplies the New element *the fourth [survey]*, while the presupposed information identifies the Old eventuality of *getting the flat*:

- (5-94) JD: Well, we lost three surveys before we got our flat
PT: And it was on the fourth you got it

A variation on this structure of cleft is the case in which the eventuality described in the presupposition is not Old, but inferrably related, to previous context. Such a case occurs in (5-95), where the presupposed eventuality is related to previous context by the inferrable relation of NEED, described in the previous section:

- (5-95) kv: I don't like going to popmobility. It's too energetic. It's more
the bendy, flexy stuff I need.

As Geluykens [1984:35] points out, *it*-cleft heads in these cases are usually very short in length. In the data in sub-corpus B, the heads of clefts of this type were usually only

one word in length, with three words as maximum.

The second category of *it*-cleft is one in which the information structure is reversed: in this case, the cleft head carries Old information, and the presupposition specifies a New or Inferrable eventuality. An Old-New case appears in (5-96):

- (5-96) B: You see one has often seen this happen
A: it's this sort of thing that makes an absolute fool of Mallet

(5-97) shows a case in which the head is Old, but the presupposition Inferrable by the relation of SOURCE, described above. JD and MM are discussing the regional origins of a particular syntactic construction:

- (5-97) JD: All you have to do is go to Leicestershire
MM: Or Lancashire. It was in Lancashire that I heard it.

The *it*-cleft therefore has two distinctive patterns, one in which a small amount of New information appears before the presupposed, Old or Inferential information; the other in which the cleft head contains Old material apparently acting as a connective, and the the New information in the presupposition. The first type is outnumbered in the corpus by the second type by 14 instances to 36, even though this structure is the one often considered to the 'default' for *it*-clefts, as we noted in chapter 2.

The Information Structure of *Wh*-clefts

Wh-clefts are much more constrained in their information structure than the *it*-cleft. As we saw in the analysis of presupposition in section 5.3, the presupposed information in *wh*-clefts must be Inferrable, or it must be Old. Although the latter case is, as we noted in chapter 2, that generally expected for *wh*-cleft, of the 50 examples examined in the data only three Old-presupposition-New-head cases appeared. The remaining cases featured Inferrable presuppositional relations.

The information specified by the *wh*-cleft head is similarly restricted, since it can only be New. We can show this by means of both anaphoric and non-anaphoric examples, as in (5-98a) and (5-98b) respectively:

- (5-98) A: I want to go to the cinema.
a B: Oh good. #What I want to do is that.
b B: Oh good. #What I want to do is go to the cinema.

Compare B's responses, however, with the reverse *wh*-cleft in (5-99):

- (5-99) B: Oh good. That's what I want to do.

In addition, it was noted in section 5.3 that the *wh*-cleft is the only one of the three cleft types that allows New, non-presupposed information to be placed in rightmost position in the sentence. We will return to this point in the discussion of clefts and sentence processing in section 5.5.

The *wh*-cleft as a whole, then, represents the possibility of prefacing large amounts of New information with a presupposition that is closely related to previous discourse. Note, too, that there is no obligation for the New information to be contrastive, which (as we noted in the introduction to this section) appears to be the case for the short head elements of *it*-clefts and reverse *wh*-clefts when they contain New information.

The Information Structure of Reverse *Wh*-clefts

The discussion of the presuppositional relations of reverse *wh*-clefts showed that New, Inferrably Related, or Old information could be carried in the presupposed part. In the entire corpus, however, only one category of head appears, namely Old. Information in reverse *wh*-cleft heads may be contrastive, but even in those cases it is Old on our definition, being either anaphoric or consisting of a definite referring expression. Very rarely, the head element is inferentially related to previous context, for example by means of a set-subset or part-whole relation. Examples of the anaphoric and the inferentially-related case appear in (5-100a) and (5-100b) respectively:

- (5-100) a A: I thought you were going to Spain.
B: Yes, that's what I told Mimi.
b A: How are you with things like cabbage and cauliflower?
B: Vegetables are what I should eat most of, so they're fine.

Ward [1985] has suggested an analysis of topicalised constructions in which an anaphoric or inferential relation can always be traced between the fronted element and the pre-

ceding context. From the reverse *wh*-cleft data analysed for this thesis, it appears that the analogy between topicalisation and the reverse *wh*-cleft would be an interesting avenue for further research, since it seems likely that similar constraints exist for reverse *wh*-clefts. For the moment, however, we will be content with the conclusion that the information borne by the head element is Old, since this accounts for all but one of the examples we have come across, the exception being (5-100b) above. It may be recalled from chapter 3, in addition, that reverse *wh*-clefts featured anaphoric NPs as head constituent in 145 out of 160 examples. As was the case in *it*-clefts, head constituents of reverse *wh*-clefts are very short in length, most usually a single syllable.

An example of the combination of Old head and New presupposed information appears in (5-101):

(5-101) JD: I've been given a really nice Habitat holepunch for Christmas.

It's matt black.

MM: Can't you go and exchange it for a stapler? Because that's what I really need.

Note that the reverse *wh*-cleft provides the best opportunity for an anaphoric head element to appear near to its antecedent, in the manner encouraged by the serial ordering principles discussed in section 5.2. The next most suitable cleft type for this purpose, the *it*-cleft, forces the interpolation of *it is* before an anaphoric head. In addition, as we noted in chapter 3, some anaphoric items, notably *it* and *that*, are marginal or unacceptable as *it*-cleft heads, reducing still further the suitability of *it*-clefts for carrying anaphoric heads.

(5-102) and (5-103) contain examples of reverse *wh*-clefts with the combinations of Old head-Inferably related presupposition, and Old head-Old presupposition, respectively. In the Old-Inferable case, the inferable relation is the relation of ORIGIN defined in section 5.3. Note also that in the Old-Old case, the conjunction of the presupposition and the information conveyed by the head are still New, as stipulated by the Informativeness requirement, defined above:

(5-102) JD: They call it a 'Bain Marie', which means 'Mary's bath'. I don't know why. Maybe she was who invented it.

relations, while New information always appeared in head position. Reverse *wh*-clefts were exactly the opposite, having obligatory Old heads, which could also be contrastive, and New or Inferrable presupposed information. *Wh*-clefts and reverse *wh*-clefts can also be represented diagrammatically in the same way as the *it*-cleft, in figures 5.8 and 5.9 respectively.

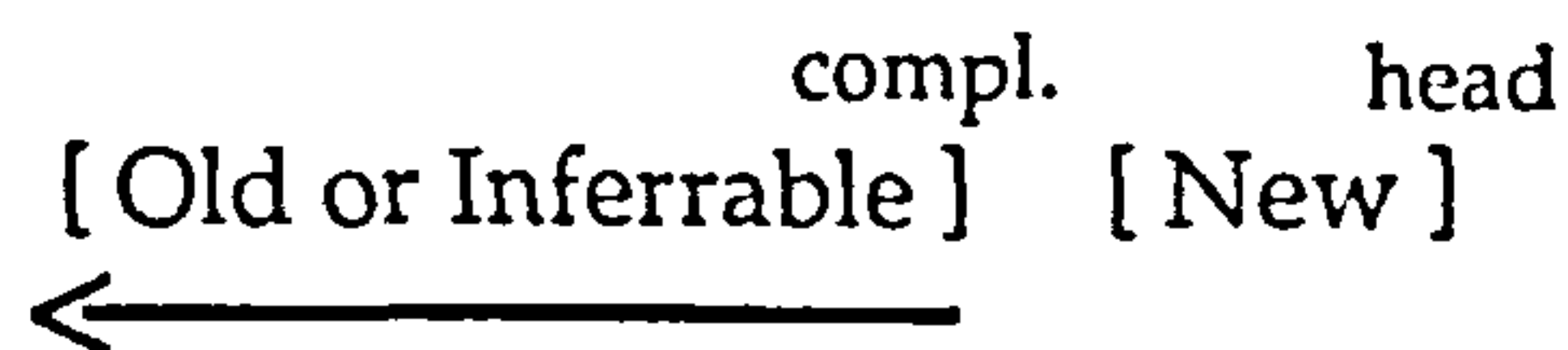


Figure 5.8: Information structure predominant in *wh*-clefts

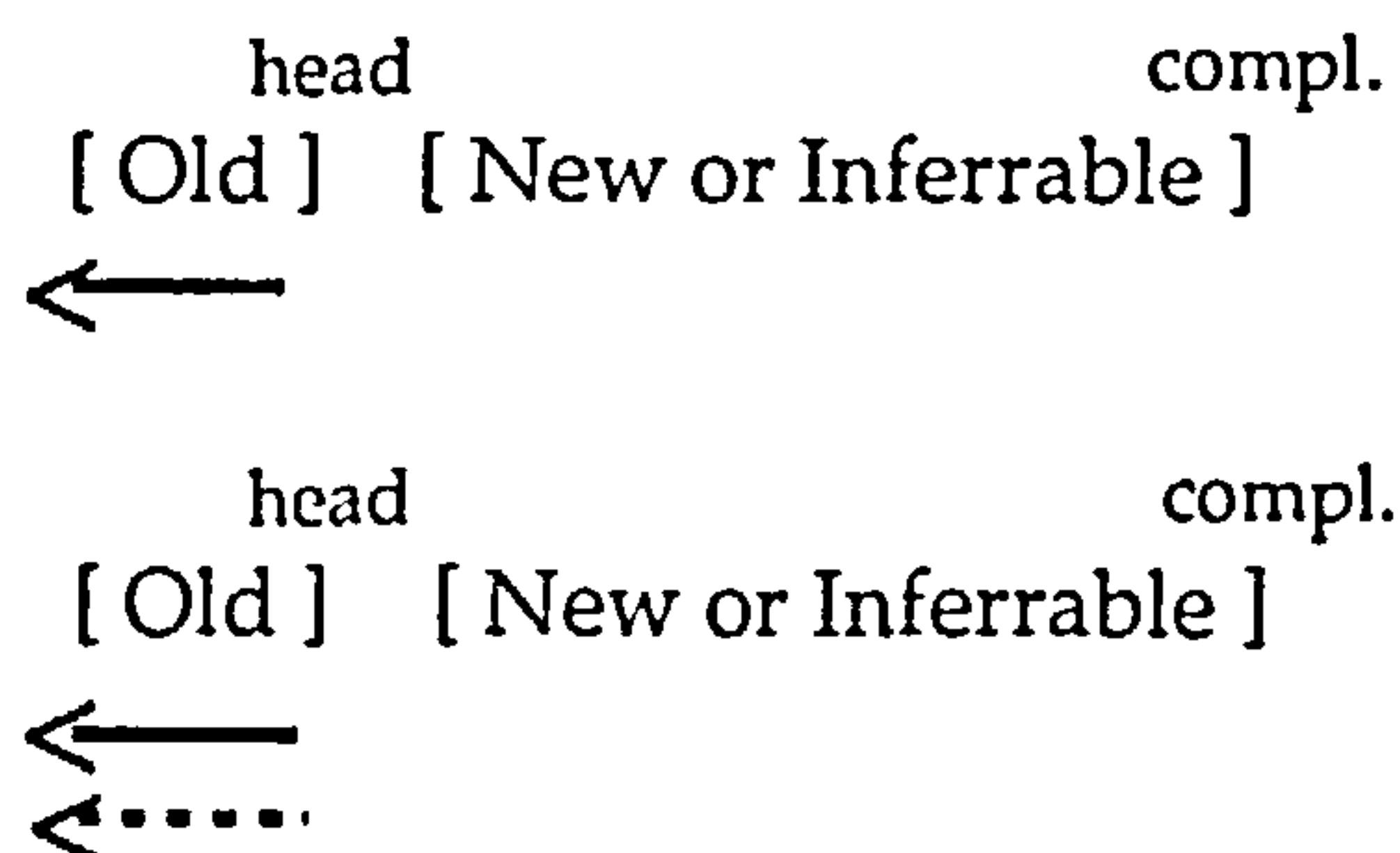


Figure 5.9: Information structure predominant in reverse *wh*-clefts

Note that of all the combinations of Old, Inferrable, or New information with presupposed or non-presupposed information, only one seems to be duplicated: the information structure shared between *it*-clefts and reverse *wh*-clefts with Old heads and New presuppositions. One distinction between the two was suggested above, however, in terms of the relationships between the anaphoric elements that each might have as head and their antecedents: those appearing in reverse *wh*-cleft heads would inevitably be closer. In addition, as we saw in chapter 3, syntactic distinctions exist between the two: how these might come into play in the choice between them in context will be discussed in chapter 6.

5.5 Information Structure and Presupposition: A Synthesis

In the discussion above, we observed two apparently disparate systems at work, as follows:

1. The syntactic marking of presupposition
2. The distribution of Old and New information in the various parts of the cleft.

In this section, we attempt to draw together the roles of these two systems in providing pragmatic markings for the messages conveyed by clefts.

5.5.1 The Relative Roles of Presupposition and Information Structure

The examination of presupposition and information structure above revealed that a wide variety of information types existed apparently independently of whether the part of the sentence in which they appeared was the part syntactically indicated to contain presupposed material. That is, information structure could not be shown to correlate reliably with what the syntactic structure of the sentence seems to signal: instead, the two systems seemed to have separate roles. These roles can be summarised as follows:

The Role of Information Structure: to reflect the current state of the discourse context, a function also served by sentence accent.

The Role of Presupposition: to mark the requirement in context for an antecedent: that is, to prescribe the state of the context required for the sentence as a whole to be interpreted.

We can therefore see information structure and sentence accent as descriptive of the speaker's expectation with respect to the state of the discourse context, while presupposition acts as a statement of the speaker's aims with respect to that context, i.e., to state what it should be like.

These two roles, however, are not independent. Taking sentence accent as an observable sign of information structure, we can demonstrate that, if the speaker's expectations regarding the discourse context as signalled by accent appear to be incorrect, the syntactically-signalled presupposition cannot be accommodated. We can demonstrate this by means of an example in which the falling intonation and lack of accenting in the cleft complement seem to be indicating that the presupposed information conveyed is already available to the hearer. In a case such as (5-104), where the presupposition is not available in this way, the result is not accommodation of the presupposition, but an ill-formed discourse. For example, A might say to someone just back from an outing:

(5-104) #Oh, you're back. It was JOHN who phoned a moment ago.

It is as if sentence accent and information structure act as signals indicating how accommodation should take place, and in the absence of correct information of this kind prevents the hearer being able to isolate the correct strategy to adopt. When working correctly, sentence accent and information structure enable the hearer to detect the nature of the discrepancy between the actual state of the discourse, and the intended state as indicated by the presupposition, and to choose an appropriate strategy to rectify this discrepancy.

5.5.2 Why Presuppose?

While the respective roles of presupposition and information structure should now be clear, we have not answered the question of *why* speakers and writers would want to mark information as presupposed in the first place. We can suggest some answers to this question by looking at the special functions that clefts can serve in discourse.

How Cleft Presuppositions are Used in Discourse

Two particular functions we can point to depend upon the use of Old and New presupposed information respectively. Old presuppositions serve the purpose of reaching back into the existing context to locate a proposition, for either of two reasons:

1. to fill in a missing value in an existing proposition; or
2. to challenge an existing value in an existing proposition.

Examples of both functions, which we can term VALUE-FIXING and CHALLENGING respectively, are as follows. An example of value-fixing occurs in (5-105):

- (5-105) A: Who made the beds?
B: It was John ϕ .

A cleft challenging a value appearing in a previous proposition is given in (5-106):

- (5-106) A: John made the beds.
B: No, it was Mary ϕ .

New presuppositions seem to serve the function simply of filling in background information about an element referred to by the cleft head. This information appears to be background not in the sense that it is less 'salient' than information appearing elsewhere in the cleft, but because it describes information away from the main stream of the narrative. An example from the written corpus illustrates this particularly well:

- (5-107) Mr. Butler, the Home Secretary, has decided to meet head-on the biggest challenge to Government authority yet presented by the Ban-the-Bomb demonstrators. Police leave has been cancelled and secret plans prepared to deal with the mass sit-down rally planned for Sunday in Parliament-square by the Committee of 100, the anti-nuclear arms group. *It was Mr. Butler who authorised action which ended yesterday in 32 members of the Committee of 100 being imprisoned for inciting a breach of the peace.* The committee's president 89-year-old Earl Russell and his 61-year-old wife were each jailed for a week.

In this case, the presupposition provides information about the main protagonist of the discourse that is relevant to the narrative, but does not constitute an event in that narrative. The observation that cleft presuppositions convey 'background' in this sense is clearly related to the constraint observed earlier on the presupposing of propositions rhetorically related to the previous discourse by simple narrative subsequence. It seems that this consideration of 'backgroundness' is bound up with issues of tense, discourse structure, and rhetorical structure, but further research is clearly necessary to establish its exact nature.

Inferable presuppositions seem to serve a dual function. Often, they can serve as an *indirect* challenge or value-fix, bringing the proposition to be challenged, or for which a value is to be supplied, into play by virtue of an inferable relation with that proposition. An example of indirect value-fixing is as follows:

- (5-108) A: Who made the beds?
B: Well, it was John who left last.

Assuming that B is being co-operative, A is expected to infer the relation between her question and B's answer, in which it appears to be suggested that John made the beds.

An indirect challenge appears in (5-109):

- (5-109) A: This coffee you made is really yukky.
B: It was you who put the milk in.

Here, B is challenging an implication of A's remark, namely that B, being responsible for making the coffee, was also responsible for adding the milk.

Although a great deal of research has to be done before the true discourse function of cleft presupposition emerges, the types of example isolated above suggest that presuppositions serve to manipulate context to achieve effects that are distinct from simple assertion. It is as if presuppositions are closer to *metalinguistic* devices than to ordinary assertions; in that, like performative verbs in Speech Act Theory, they cause the state of the world—in this case, the discourse world—to change.

While little is currently known about the discourse functions of presuppositions beyond the fact that it occurs, more is available in terms of how presuppositions are processed in discourse. In the following section, we examine a model of human sentence processing as an analogy for the model of presupposition and its relation with context suggested so far. In particular, the model provides some interesting parallels for the strategies suggested so far for incorporating cleft presuppositions into context. Of course, it is important to note that the analogy being drawn here is not intended to claim psychological plausibility for the formal mechanisms used in this and the previous chapter for representing the relationships between clefts and context, such as DRSS and the use of axioms for inference.

Presuppositions in a Psychological Model of Sentence Processing

The psychological model of sentence processing referred to above is that described by Clark and Haviland [1977]. We can summarise the model briefly as follows.

On Clark and Haviland's account, one of the first tasks of a hearer or reader in processing a unit of language is to identify which part of the message is that which should, on the basis of the cues given by the speaker, have an antecedent in the hearer's memory. Clark and Haviland refer to this identification task as the application of the GIVEN-NEW STRATEGY, in which hearers and readers identify Given information as that which has a unique antecedent, the location of which is the address or location in memory at which the incoming information is to be stored. Clark and Haviland's claim is that hearers or readers encountering a device such as a cleft are able from the form of the cleft to recognise information for which an antecedent should be present in memory, and go on to identify antecedents for that information.

In their suggestion for clefts, Clark and Haviland appear to equate Given information with the cleft presupposition. Even though we have seen that this is incorrect on the Hallidayan notion of Given, it is still plausible to suggest that the information identified by hearers as requiring an antecedent is the presupposed information, analogous to the location or formation of antecedents for presuppositions in the DRT model discussed in chapter 4.

As we have seen, however, the information in cleft presuppositions is in the majority of cases not information for which an antecedent is already available in context. Clark and Haviland's model, however, allows for this, in that it permits information to be presented as requiring an antecedent when no antecedent is in fact available. The cleft construction can be treated in Clark and Haviland's model as a construction that uses conventional syntactic signals to indicate that the presupposed information should be treated by the hearer or reader as if it had such an antecedent, by forming one for it at the time of processing—in formal terms, by accommodating the antecedent.

Clark and Haviland's model suggests a psychological correlate to this process of accommodation. On their account, if the hearer or reader can find no direct antecedent for

presupposed information at the time of processing, three compensatory strategies are available. The three strategies are, respectively, BRIDGING, ADDITION and RESTRUCTURING. The first two of these, we would suggest, constitute a specialisation of the process of accommodation in the ways already suggested in this chapter, being analogous with the treatment of INFERRABLE and NEW presuppositions respectively.

The following definitions of the three strategies are adapted from Clark and Haviland [1977:6ff]:

Bridging: When a listener or reader cannot find a direct antecedent for information marked as requiring one, he or she may be able to form an INDIRECT ANTECEDENT by building an inferential BRIDGE from something he or she already knows.

Addition: Sometimes it is impossible to form an inferential bridge between the incoming information marked as requiring an antecedent and an actual antecedent in memory. In this case, the hearer must ADD to memory something that will act as an antecedent, possibly newly constructed from the content of the message he or she is trying to process.

Restructuring: Restructuring is used as a last resort, when the material that is marked as antecedent-requiring does not seem to be the correct material to use as any kind of basis for a coherent relationship. In cases such as this, the hearer can re-interpret the markings of the incoming message until an interpretation is arrived at which enables either the location of an antecedent, or bridging or addition, to take place.

While Bridging can be correlated with the treatment of Inferrable presuppositions, and Addition with the treatment of New presuppositions, we can assume Restructuring to be simply an 'error condition'.

If we adopt the assumption that the cleft presupposition is the processing cue for finding addresses in memory, which is consequently processed first, this would provide a general explanation for the observation made in section 5.4 that clefts whose heads appear prior to the presupposed information do not carry a large amount of information in head position. On the application of Clark and Haviland's model suggested here, the

information carried in the cleft head would have to be held in short-term memory until the presupposition had been processed. The processing overhead associated with this would increase in proportion to the amount of information that needed to be stored in this way: presumably, the longer the cleft head, the greater that processing load.

Of course, without further experimentation on clefts of all three kinds, the analogy above cannot be considered to be more than a plausible suggestion. In particular, it needs to be established what effects a variety of information structures and accentual patterns has in the processing of clefts: whether, for example, a head would be processed first if it were anaphoric, and not the presupposition as suggested above.

5.6 Pragmatic Aspects of Cleft Constructions: Summary

In this chapter we have addressed several interdependent areas surrounding the use of clefts in context. In particular, we have been careful to point out factors that differentiate the three cleft types from one another, and, where possible, factors that differentiate clefts from non-cleft sentences. The main points of the chapter can be summarised as follows:

- Clefts have a wide range of accentual patterns, of a far greater variety than has generally been associated with them in the literature.
- While the information-ordering principles based on 'Given before New' or on SVO ordering cannot be upheld, the psychological notion of REFERENTIAL CONTINUITY is relevant in the description of clefts, particularly the choice between them.
- Apparently, propositions standing in the rhetorical relationship of SIMPLE NARRATIVE SUBSEQUENCE cannot appear as the presuppositions of clefts.
- The relationships of cleft presuppositions to discourse context can be captured in terms of three categories, OLD, INFERRABLE and NEW. Differing constraints operate on clefts of all three types in relation to which of these kinds of presupposition they can carry.
- Although the hypothesis that clefts serve to differentiate New information from Given could not be upheld, two generalisations could be made: one, that clefts

obligatorily communicate a novel connection between the set of elements specified by the cleft head, and the existentially-quantified variable in the cleft presupposition, and two; that no cleft can be composed entirely of New information.

- It is also possible to demonstrate distinct and typical information structures for each of the three cleft types.
- The roles of presupposition and information structure, though distinct, are not independent. Presupposition serves to identify the required state of the context in order for the interpretation of the sentence to be completed; information structure relates to the current expected state of that context. Information structure enables the interpreter to detect the discrepancy between the required and current state and to select the appropriate strategy for closing the gap.
- Analogies were drawn between the formal model of cleft presupposition discussed in this and the previous chapter, and the model for human sentence processing described by Clark and Haviland [1977]. It was suggested in particular that the relationship between presuppositions and context were formed by accommodation, namely the Inferred and New cases, were analogous to Clark and Haviland's processing strategies of Bridging and Addition.

In the next (the final) chapter of this thesis, we suggest how the conclusions summarised above might be put to use in a decision procedure for syntactic choice.

Chapter 6

Conclusions and Further Research

So far in this thesis we have looked at three main aspects of cleft constructions: their syntax, their semantics, and the pragmatic factors surrounding their use in discourse. In this chapter, we summarise the main claims made in previous chapters with respect to each of these areas, and draw some general conclusions about how they interact.

The factors that we have concentrated on in this research on clefts fall into two broad groups, as follows:

- factors that differentiate clefts from other sentence types; and
- factors that differentiate one type of cleft from another.

Factors of the first type have been predominantly semantic in nature, and relate to the function of cleft sentences as a group. Factors of the second type have been pointed out mainly at the level of syntax and pragmatics, emphasising how the three types of cleft in English serve to provide a variety of options with respect to the content that can be conveyed, the pragmatic markings that can be communicated along with that content, and the way in which the message as a whole fits into context. In the summary of the claims of this thesis given in section 6.1 below, these two groups of factors will be brought out in particular. They are further emphasised in section 6.2, where the conclusions reached in the thesis are synthesised into an outline for a decision procedure

for choosing between a cleft and a pragmatically-unmarked syntactic construction, such as a declarative, for the purposes of conveying a given message in a given context.

Finally, as is common to all research, there are many questions that remain to be answered if a complete understanding of the form and function of cleft constructions is to be fully appreciated. Accordingly, in section 6.3, we point out some unsolved problems that have arisen during the course of the exposition, as well as suggesting some interesting related directions for future research.

6.1 The Claims Made in this Thesis

In chapter 2 of this thesis, we examined previous research on clefts in discourse, concentrating on claims to the effect that clefts mark particular kinds of information status on the information they carry, such as GIVEN, FOCUS, and NEW. We found that most claims of this nature were inaccurate or insufficiently general due to a lack of attention to the variety of the data, and in particular due to the assumption that factors such as information structure could straightforwardly be associated with syntactic structure. It was demonstrated, for example, that syntactic structure did not reliably mirror either Given-New structure, or associate with particular accentual patterns.

Apart from the variety inherent in the data, we were able to show in particular:

- No definition of FOCUS so far suggested in relation to clefts can reliably be associated with clefts of all kinds, which casts considerable doubt on the popular assumption that clefts are ‘focusing’ constructions.
- On a related topic, the assumption that cleft heads are a reliable source of antecedents for pronominal anaphora, built into some prominent computational theories for anaphora resolution, does not appear to be supported by the data.
- Explanations of the existence of clefts couched in terms of the serial ordering of information or syntactic constituents were shown to be ill-founded, on the basis of a variety of counterexamples.

Most striking about the research reviewed, however, was its uniform tendency to suggest

explanations for the existence of clefts in terms of a single layer of linguistic description. Examples of this are the theory that clefts serve to indicate information structure, or that they provide options for a particular serial ordering of information. Such theories, which tend to reside particularly at the level of pragmatics, neglect the fact that cleft constructions were *additionally* differentiated from other sentence types at other levels, notably at the levels of syntax and semantics. For example, most of the pragmatics research reviewed entirely neglected the fact that clefts were presuppositional constructions, thereby losing themselves the often saving explanation that, even if a cleft was not the only syntactic construction that could offer, for example, a particular serial arrangement of information, it may still be the only *presuppositional* construction that could do so, and thereby still be an attractive option for the speaker.

On the basis of this objection, it was argued that a multi-layer approach to the explanation of marked syntactic structures was required, in which it was possible to combine claims about the pragmatic functions of clefts with observations regarding their differentiation from other syntactic types at levels of description other than pragmatics. Using this as a starting point, the thesis set out to examine what distinctions between clefts and other sentence types, and between the three types of clefts themselves, resided at each level, and how these distinctions could be considered as part of a system of factors determining the unique role of each type of cleft in discourse.

Accordingly, in chapter 3 we set out to look at how the three types of cleft could be characterised at the level of syntactic structure. GPSG rules were formulated for each type of cleft (or, in the case of the *it*-cleft, existing rules were available from within the GPSG framework). We showed how the various syntactic characteristics of the three cleft types could be captured by these rules, and how the clefts, or syntactic constituents within them, could be differentiated from other superficially similar constructions. In particular, we examined the types of constituent that could appear as head in each type of cleft, not only for the completeness of the analysis, but also because the distinctions in this direction were likely to be at the root of distinctions in the content it was possible to convey using each type of cleft. Apart from observing distinctions in the possibility of anaphoric elements appearing as head of each type of cleft, the potential for other syntactic constructions to appear in that position was given as a table, repeated here in figure 6.1.

	<i>it</i> -clefts	<i>wh</i> -clefts	reverse <i>wh</i> -clefts
NP	✓	✓	✓
Finite VP	*	*	*
Purposive <i>to</i> -infinitive	✓	✓	✓
Non-purposive <i>to</i> -infinitive	*	✓	✓
- <i>ing</i> as VP complement	*	✓	✓
Subject <i>ing</i> participle	✓	✓	✓
Passive participle	✓	✓	✓
Bare infinitive	*	✓	?
NP + <i>not until</i>	✓	*	*
XP + <i>not only/just</i>	✓	✓	*
Universally-quantified NP	*	✓	*
Negative polarity NP	*	✓	*
AP	? *	✓ (most)	? *
- <i>ly</i> ADV	*	*	✓
ADV P	✓	✓	✓
Sentential ADV	*	*	*
s complement of factive v	✓	✓	✓
s complement of non-factive v	*	✓	?
PP	✓	✓	✓

Figure 6.1: Possible Head Constituents by Cleft Type

In chapter 3, we also examined the *actual* occurrence of the strings predicted by the grammar in the corpus of data. It was observed that only a subset of the possible strings appeared with any regularity, a fact that was attributed to pragmatic factors.

In chapter 4, we turned to what was seen to be the core of the distinction between clefts and other sentence types: their semantics. We differentiated between clefts and other sentence types on the basis that clefts of all three types were PRESUPPOSITIONAL constructions, on a conventional definition of the notion, whose content could be represented as a presupposition containing an existentially-quantified variable, and an assertion which served to link that variable with a value. On the basis of van der Sandt's [1988] suggestion that presuppositions could be treated as anaphors within the framework of Discourse Representation Theory (DRT), we worked out the detail of such an account, concentrating in particular on how cases in which the presupposed information was not already present in the context (cases requiring ACCOMMODATION, after Lewis [1979]) could be dealt with in the framework. The mechanism offered in DRT for dealing with these cases, and with cleft presuppositions in general, was straightforward,

relying only on the introduction of an eventuality marker for each predicative VP in the sentence, and on a distinction, made primarily in the interests of clarity, between the PRECONDITIONS on the interpretation of a sentence and the POSTCONDITIONS of having interpreted it.

In addition to the fact that clefts are presuppositional constructions, a second important aspect of the meaning of clefts that is not associated with the corresponding declarative sentences is that the element (or set of elements, on our formulation) specified by the cleft head was UNIQUE or MAXIMAL with respect to the predicate in the presupposition. We captured this by a condition MAX in the DRT formalism, the derivation of which condition was based on an analogy between the behaviour of the existentially-quantified variable concerned and ordinary definite referring expressions, as analysed by Hawkins [1978].

In chapter 5, we turned to the pragmatic level of describing clefts, examining some separate but related aspects of their use in context. We saw first of all the range of accentual patterns found in a corpus of 150 examples, the data serving to further substantiate claims made in chapter 2 regarding the variety in the data. Second, we went on to show how the psychological principle of the preservation of REFERENTIAL CONTINUITY provided an information-ordering principle relevant for the choice between cleft types that was able to account for data that principles such as Given-New ordering (reviewed in chapter 2) failed to explain. We then went on to examine the range of relations that could exist between cleft presuppositions and preceding context, providing a detailed analysis to substantiate the claim made in chapter 4 to the effect that the information in cleft presuppositions did not always have to be present in the context. In fact, we defined and gave examples for three types of presuppositional, OLD, INFERRABLE and NEW, and showed the differences between the three types of cleft with respect to their capacity for exploiting the three kinds of relation. In particular, we saw that *wh*-clefts could not carry New presuppositions, while *it*-clefts could not convey presuppositions related to context by Subsumption, a subdivision of the Inferrable category. It was also demonstrated that a general constraint existed on cleft presupposition in that propositions having a particular rhetorical relation to preceding discourse, which we termed SIMPLE NARRATIVE SUBSEQUENCE, could not be presupposed at all. It was suggested that this was one of a possible range of rhetorical and discourse-structure constraints

on the acceptability of propositions as cleft presuppositions.

Turning then to the relationship of clefts as a whole to the preceding context, we used our analysis of presuppositional types as part of a framework for analysing information structure in general, that study being based on a linearised DRT representation of the sentences involved. We saw further evidence against the claim, originally reviewed in chapter 2, that information structure correlates reliably with syntactic structure, and also observed two general features regarding the information structure of clefts. The first of these was that a cleft had at the very least to contribute an informative connection between the set of elements specified by the head constituent and the existentially-quantified variable in the presupposition conveyed by the complement—that is, that this connection must not already be entailed by the context. We termed this the MINIMAL INFORMATIVENESS REQUIREMENT. We suggested that this requirement exposed a further aspect of the discourse function of clefts, namely that of drawing attention to the asserted content, that assertion being the connection between the cleft head element and the variable in the presupposition. Secondly, we observed that clefts also had to be at least MINIMALLY COHERENT with context, in that they had to contain at least one non-New element.

Finally, we examined the information structure of all three cleft types, and observed the various distinctions between them. The most notable of these was that *wh*-clefts could not convey Old information in cleft head position, while reverse *wh*-clefts could not convey New information in this way.

In the final section in the chapter on cleft pragmatics, we gave a synthesis of the relative roles of syntactically-signalled presupposition and information structure. It was claimed that presupposition indicated the *required* state of the context for the cleft content to be integrated, while information structure, and relatedly sentence accent, indicated the *actual* state of that context as the speaker assumed it to be. We then examined the residual question of why presupposition takes place at all, suggesting that speakers used it to reach back into the discourse context to locate available propositions for the purposes of challenging them or adding to them, or to provide background information about a current element in the discourse.

In conclusion, we looked at a psychological model of sentence processing that offered

interesting analogies to the formal model of cleft interpretation developed in chapters 4 and 5. We suggested that the relationships of presuppositions to context found in the corpus were reflected in the processing strategies of BRIDGING and ADDITION suggested by Clark and Haviland [1977]. On the basis of their model, it was suggested that cleft presuppositions were processed first, which accounted for features of the information structure of the data that were observed earlier—notably that clefts whose presupposed information was presented after the cleft head had significantly shorter head constituents than those in which this was not the case, presumably to avoid processing load on short-term memory.

In section 6.2, we give a brief overview of how these various features might fit together in an outline procedure for syntactic choice.

6.2 Choosing Clefts in Context

In this section, we take a brief look at how the conclusions reached in the thesis might fit together in a decision procedure for the computer generation of appropriate clefts in context. For this purpose we can assume a state of affairs in which the discourse function and content of an utterance are planned, but how that content and function is to be given a surface realisation is not (cf. the ‘what to say vs. how to say it’ distinction, discussed in Thompson [1977].)

As suggested above, we might see the generation of clefts as essentially a two-stage process: the choice of whether or not to use a cleft at all, and the choice of which type of cleft to use.

To Cleft or not to Cleft?

Given the content of an utterance (which we can suppose to be represented as a proposition) and a context, the first decision that would plausibly be made is the decision regarding whether or not to use a cleft. We can imagine for the purposes of the first process that there exist two kinds of sentences in the world: clefts and declaratives. (Of course, many other sentence types exist, but complex constraints operate on their

use that are outside the scope of this thesis, and I do not therefore intend to address them here.) The first thing to decide is whether it is desirable to give the content of the message the particular characteristics that we claimed in this thesis to be the function of the cleft.

We can briefly summarise these claims for the functions of clefts differentiating them from declarative sentences, as follows:

- Clefts serve to draw attention to the asserted relationship between the (possibly singleton) set of elements specified by the cleft head and the existentially-quantified variable in the presupposition.
- Clefts are presuppositional constructions.
- Clefts convey the fact that the set of elements specified by the cleft head are unique in the context with respect to their ability to provide an instantiation of that variable.
- Clefts require at least one discourse object in the domain of things referred to by the cleft to be currently available in the context. On this basis, they serve to highlight cohesiveness explicitly.

To decide whether these markings are required, the corresponding questions to ask might be those below:

- Is it required to draw attention to the material that would be asserted? Does that assertion satisfy the minimal informativeness requirement, namely that it be New in the discourse?
- Is it desirable to mark part of the message as presupposed, taking into account the function of presupposition as discussed in chapter 5?
- Given that presuppositional marking on part of the message, does the residue of the message—the part that would be arranged as cleft head—satisfy the MAXIMALITY criterion with respect to the predicate in the presupposition?
- Is the message to be conveyed such that an explicit cohesive link with the prior context can be maintained?

If the answer to all the above conditions is affirmative, it is likely that a cleft is a good candidate for conveying the message. The second stage of the procedure can then be invoked: deciding which cleft to use.

Which Cleft?

This part of the procedure is made up of a large number of interacting constraints, and it is not always easy to see the best means of prioritising them. The ordering suggested here is intended merely to be indicative of how they might interact. It is unlikely that they can be satisfied all at once, and given the eventuality of one or more being violated, it is up to the researcher to decide how serious a violation this is: can it be tolerated, or does it mean that the particular cleft type must be completely ruled out? Are there interactions between the constraints suggesting that particular *combinations* of violations are worse than others? In a fully-specified algorithm, it is likely that a system of weightings and penalties might be of use in deciding on the gravity of particular violations or combinations of such violations.

On the basis on the first part of the procedure, we can expect the information to reach this second part to be a representation of the content of the cleft such as we discussed in chapters 4 and 5, namely a specification of which information is presupposed, and which asserted. This specification needs then to be mapped into a range of possible syntactic realisations. This might be achieved by concentrating in particular on the possible syntactic realisations of the semantic object that provides the value with which the variable in the presupposition is to be instantiated—that is, the content that is to be realised as the head of the cleft. This would be done taking into account the conclusions on acceptable head types reached in chapter 3.

Of course, once the realisation of this is decided upon, other features of the syntax, in particular the dependencies described between elements in the structure (both gap-filler dependencies and those dependencies captured by feature co-occurrence restrictions) will serve to determine the syntactic structure of the remainder of the cleft. At this stage, therefore, we can expect to have a range of possible realisations of the content, perhaps more than one of each type of cleft. At this stage, constraints at various levels can

be brought into play to decide between the sentences, successively narrowing down the possible alternatives.

While no claims are made here regarding the priority of the various constraints, it is suggested on the basis of the conclusions reached in this thesis that the following factors are relevant in reducing the range of possible alternatives:

- If the cleft head has to be anaphoric, some cleft types may be ruled out because of the unacceptability of these elements as head.
- Related to the above, some cleft types may be ruled out because the relationship between the content of the cleft and the preceding context, which can be described as the INFORMATION STRUCTURE of the cleft, is inappropriate for the particular cleft type in the manner described in chapter 5. This may apply to the presupposed information, the information borne by the cleft head, or the combination of the two.
- The space of alternatives can further be narrowed down by the observance of constraints on processing: the chosen cleft should be maximise the degree of referentially continuity between its content and the surrounding discourse, and should minimise postponement of the presupposition.

If there is more than one sentence remaining at the end of this process, assuming that the constraints and their interaction have been set up correctly, the resulting sentences should be interchangeable in the context with no loss of acceptability, in the manner described in chapter 1. If no sentences remain, this suggests that a declarative (or other sentence type, if similar decision procedures can be formulated for them) is the appropriate choice.

6.3 Directions for Further Research

As noted at the beginning of this chapter, many questions remain to be answered concerning the structure and function of clefts in discourse. Some of these arise as unsolved problems in the current research, others are simply interesting avenues for future ex-

ploration. The issues will be treated under the various headings of the chapters in the thesis, as follows:

- Syntactic Issues
- Semantic Issues
- Pragmatic and Processing Issues

We will look at each one in turn.

Syntactic Issues

In the discussion of cleft syntax in chapter 3, we saw that cleft constructions of various kinds had partially distinct constraints on what constituents could appear as cleft heads. We did not, however, give rules for each different case, and could not suggest *why* the cleft types should be distinct in this way, nor why they should be distinct from other constructions—such as topicalisation—in which a similar kind of extraction seems to be taking place. Both rules and explanations remain to be formulated.

Although we were able to differentiate some superficially-similar constructions from the clefts under examination, the analysis of one particular structure still remains a mystery: are sentences such as *it's an ill wind that blows nobody any good* actually clefts, *it*-relative constructions, or something different again?

On a related note, how justified is the distinction we have made between clefts as they have been studied in this thesis and copular sentences such as the following?

- (6-1)
- a The one I am looking for is Liberty.
 - b She's the one I want.
 - c All I care about is greengages.
 - d The thing she's wearing is a kaftan.
 - e The girl in the corner is my aunt's bridge partner.

In other words, where do clefts stop, and other types of copular sentence begin?

Semantic Issues

The DRT model for cleft presupposition presented in chapter 4 could well be extended, particularly to deal with clefts in which negation is involved. What, for example, is the treatment of clefts such as *it was not only John who smiled*? How do these *not only* and *not just* clefts interact with the criterion for maximality postulated? On a related issue, how can the fact that only *it*-clefts are completely acceptable with *not only*, *not just* and *not until* be accounted for?

We worked out the details of presupposition for *it*-clefts in chapter 4, and suggested that the treatment can straightforwardly be generalised to deal with the other two cleft types. No semantic distinctions have been discovered between them during the course of the thesis except for the fact that *wh*-clefts alone can appear with quantified NPs and the complements of non-factive verbs as head. This suggests that the head of the *wh*-cleft might not have the same kind of 'definiteness' that the *it*-cleft and the reverse *wh*-cleft have, indicating in turn that a different semantic treatment might be necessary for *wh*-clefts. Are these factors evidence for this, or can they be explained more superficially?

Pragmatic and Processing Issues

Some of the most interesting areas for further investigation concern what might be called pragmatic issues. Some suggestions are as follows.

In the discussion of why presuppositions are used at all, we pointed out some discourse functions of clefts, for example VALUE-FIXING and CHALLENGING. What other discourse functions can be observed for clefts?

On a related issue, we noted that one of the effects of cleft presuppositions that contained New information was to provide background information. We also related this to the fact that some rhetorical relations, such as simple narrative subsequence, did not seem to be good subjects for presupposition. Using these facts as clues, what are the relationships between clefts and the rhetorical structure of discourse? Do they, for example, signal certain kinds of subordination relations among the propositions in discourse? How

might their effects be represented, for example, in a formal model of discourse structure such as that of Grosz and Sidner [1985]?

In the discussion of the information structure of clefts, we noted that the short head constituents of *it*-clefts and *wh*-clefts tended to result in a contrastive reading when New information was conveyed by those constituents. Taking a theory of the relationship between sentence accent, New information, and the phonological notion of 'Focus', (cf. Gussenhoven [1983], Selkirk [1980], *inter alia*) how can the fact be explained that the absolute *length* of an accented constituent seems to affect the interpretation of the focal scope of an accent? Might it be that clefts signal SYNTACTIC areas over which accents have scope?

We noted the acceptability or unacceptability of anaphoric items as head in the various cleft types as a simple fact. However, it is well known that the appropriateness of anaphoric elements is pragmatically determined. Can the constraints on anaphoric heads be formulated as a pragmatic constraint on clefts? Can distinctions in pragmatic function of the three cleft types be formulated on the basis of their differing abilities to accept anaphoric heads?

Similarly, we noted regularities in the information structure of clefts of all three types. However, what is the explanation behind these regularities? Why are the cleft types different? Might the serial ordering of information internal to the cleft, coupled with the different positions available for presupposed information, be at the root of the different possibilities that exist in the three cleft types for the siting of information? Might phonological factors be involved?

In chapter 5, we suggested some reasons as to why people presuppose. Is this the whole reason? Relatedly, why is presupposition routinely associated with particular syntactic structures?

We also suggested a rhetorical relation that seemed to prevent presupposition taking place. Is this a side-effect of presupposed information having to be 'background', as described above, or are there other rhetorical constraints on presupposition? How can they be explained?

Finally, why are examples such as the following unacceptable?

- (6-2) A: Is a yellow block on the table?
a B: #No, it's the box that a yellow block is on.
b B: #No, the box is what a yellow block is on.
c B: #No, what a yellow block is on is the box.

Is there a simple prohibition on presuppositions containing indefinites? If so, why?

In the discussion of clefts and processing, we made several suggestions that can only be assessed on the basis of psychological experimentation. In particular:

- What is presupposition for, in processing terms?
- What is the status of the claim that presuppositions are processed first? Is this affected by the information structure of the presupposition, or the information structure of the rest of the cleft? What, in general, is the psychological effect of presupposing (taking care to distinguish this from the effects of particular kinds of information structure)?
- We have claimed that one of the functions of clefts is to highlight the connection between the element specified by the cleft head and the predicate in the presupposition. Does this mean that the connection is being made 'salient' in some way?
- On a related note, it was argued in chapter 2 that clefts are not 'focusing' constructions on any of the definitions of focus suggested so far. This was due to the fact that focus on those definitions did not correlate reliably with syntactic structure. Are there any psychological grounds for saying that focus is relevant to clefts? Experiments to date (cf. Clark and Clark [1977], Yekovich *et al.* [1985]) do not separate out the facts about accentual pattern and information structure from the observation of the effect of cleft syntax on memory, with the result that it is not clear whether it is the cleft, or the information structure, that has the observed effect. Experiments on clefts with a controlled variety of information structures would enable us to separate out the relative roles of the two.

As is often the case, the research reported herein has produced answers and questions in roughly equal proportion. In view of the fact that such questions remain, I do not

feel that the conclusions reached in this thesis should be seen as a full stop, rather as a brief pause in the dialogue on cleft constructions.

References

- Akmajian, A. [1970] On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry*, 1, 146–168.
- Allwood, J. [1972] Negation and the strength of presupposition. Logical Grammar Reports No. 2, Department of Linguistics, University of Goteborg, 1972.
- Atlas, J. D. and Levinson, S. C. [1981] It-Clefts, Informativeness, and Logical Form. Chapter 1 in Cole, P. (ed.) *Radical Pragmatics*, pp1–59. New York: Academic Press.
- Bach, E. [1986] The algebra of events. *Linguistics and Philosophy*, 9, 5–16.
- Bolinger, D. L. [1961] Contrastive Accent and Contrastive Stress. *Language*, 37, 83–96.
- Bresnan, J. and Grimshaw, J. [1978] The Syntax of Free Relatives in English. *Linguistic Inquiry*, 9, 331–391.
- Carter, D. [1987] *Interpreting Anaphors in Natural Language Texts*. Chichester: Ellis Horwood.
- Chafe, W. [1974] Language and Consciousness. *Language*, 111–133.
- Chafe, W. [1976] Givenness, contrastiveness, definiteness, subjects, topics and points of view. In Li, C. (ed.) *Subject and Topic*, pp25–56. New York: Academic Press.
- Chierchia, G. and Turner, R. [1988] Semantics and Property Theory. *Linguistics and Philosophy*, 11, 261–302.
- Chomsky, N. [1970] Deep structure, surface structure and semantic interpretation. In Jakobson, R. (ed.) *Studies in Oriental and General Linguistics*, pp52–91. Tokyo.
- Chomsky, N. [1971] Deep structure, surface structure, and semantic interpretation. In Steinberg, D. D. and Jakobovits, L. A. (eds.) *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology*, pp183–216. Cambridge: Cambridge University Press.
- Clark, H. and Clark, E. [1977] *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovitch.
- Clark, H. H. and Haviland, S. E. [1977] Comprehension and the given-new

- contract. In Freedle, R. O. (ed.) *Discourse Production and Comprehension*, Volume 1, pp1-40. Norwood, N.J.: Ablex.
- Couper-Kuhlen, E. [1986] *An Introduction to English Prosody*. London: Edward Arnold.
- Creider, C. [1979] On the explanation of transformations. In Givon, T. (ed.) *Syntax and Semantics*, Volume 12: *Discourse and Syntax*, pp3-22. New York: Academic Press.
- Cruttenden, A. [1986] *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. [1969] *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- Daneš, F. [1974] *Papers on Functional Sentence Perspective*. Prague: Academia.
- Delahunty, G. [1984] The Analysis of English Cleft Sentences. *Linguistic Analysis*, 13, 63-113.
- Doron, E. [1988] The semantics of predicate nominals. *Linguistics*, 26, 281-302.
- Ehrlich, K. and Johnson-Laird, P. N. [1982] Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior*, 21, 296-306.
- Enkvist, N. [1973] *Linguistic Stylistics*. The Hague: Mouton.
- Faraci, R. [1971] On the Deep Question of Pseudo-Clefts. *English Linguistics*, 6, 48-85.
- Firbas, J. [1964] On defining the theme in functional sentence analysis. *Philologica Praguensia*, 8, 170-176.
- Firbas, J. [1966] Non-thematic subjects in contemporary English. *Travaux Linguistiques de Prague* No. 2, Prague, 1966.
- Firbas, J. [1975] On the thematic and non-thematic section of the sentence. In *Style and Text*.
- Fletcher, C. R. [1984] Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior*, 13, 487-493.
- Garnham, A., Oakhill, J. and Johnson-Laird, P. N. [1982] Referential continuity and the coherence of discourse. *Cognition*, 11, 29-46.
- Gazdar, G. [1976] Formal Pragmatics for Natural Language Implicature, Presupposition, and Logical Form. PhD Thesis, Department of Linguistics, University of Reading.
- Gazdar, G. [1979] *Pragmatics, Implicature, Presupposition, and Logical Form*. New York: Academic Press.
- Gazdar, G., Klein, E., Pullum, G. and Sag, I. [1985] *Generalized Phrase Structure Grammar*. London: Basil Blackwell.

- Geluykens, R. [1984] Focus Phenomena in English: an Empirical Investigation into Cleft and Pseudo-Cleft Sentences. Technical Report No. 36, Departement Germaanse, Universitaire Instelling Antwerpen, Antwerp, 1984.
- Givon, T. [1978] Negation in Language: Pragmatics, Function, -Ontology. In Cole, P. (ed.) *Pragmatics*, Volume 9, pp69-112. New York: Academic.
- Glatt, B. S. [1982] Defining Thematic Progressions and Their Relationship to Reader Comprehension. In Nystrand, M. (ed.) *What Writers Know: The Language, Process, and Structure of Written Discourse*, pp87-103. New York: Academic Press.
- Green, G. M. [1980] Some wherefores of English inversions. *Language*, 56, 582-602.
- Grosz, B. and Sidner, C. L. [1985] The Structures of Discourse Structure. Report No. CSLI-85-39, Center for the Study of Language and Information, 1985.
- Gundel, J. [1975] Left-dislocation and the Role of Topic-Comment in Linguistic Theory. *Ohio State Working Papers in Linguistics*, 18, 72-131.
- Gundel, J. K. [1977] Where do cleft sentences come from? *Language*, 53.
- Gussenhoven, C. [1983] Focus, Mode, and the Nucleus. *Journal of Linguistics*, 19, 377-417.
- Halliday, M. A. K. [1967] Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3, 199-244.
- Halliday, M. A. K. [1973] *Explorations in the Functions of Language*. London: Edward Arnold.
- Halliday, M. A. K. [1976] English System Networks. Chapter 9 in Kress, G. R. (ed.) *Halliday: System and function in language*, pp101-135. Oxford: Oxford University Press.
- Halliday, M. A. K. [1985] *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halvorsen, P. [1978] The Syntax and Semantics of Cleft Constructions. PhD Thesis, Linguistics, University of Texas at Austin.
- Harries, H. [1972] The Role of Cleft Sentences in the Formation of Questions and Answers: Some Evidence from German. In Peranteau, P. M., Levi, J. N. and Phares, G. C. (eds.) *Papers from the Eighth Regional Meeting of the Chicago Linguistics Society*, Chicago, April 14-16, 1972, pp124-138.
- Haviland, S. E. and Clark, H. H. [1974] What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Hawkins, J. A. [1978] *Definiteness and Indefiniteness*. London: Croom Helm.
- Heim, I. [1982] The Semantics of Definite and Indefinite Noun Phrases. PhD Thesis, University of Massachusetts. Distributed by Graduate Linguistic Student Association.

- Higgins, F. R. [1979] *The pseudo-cleft construction in English*. New York: Garland Publishing.
- Hirst, G. [1981] *Anaphora in Natural Language Understanding*. Berlin: Springer-Verlag.
- Hobbs, J. R. [1978] Resolving Pronoun References. *Lingua*, 44, 311–338.
- Hornby, P. A. [1972] The psychological subject and predicate. *Cognitive Psychology*, 3, 632–642.
- Huddleston, R. [1984] *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Jacobson, P. [1987] Phrase Structure, Grammatical Relations and Discontinuous Constituents. In Huck, G. J. and Ojeda, A. E. (eds.) *Syntax and Semantics*, Volume 20: *Discontinuous Constituents*. London: Academic Press.
- Jespersen, O. [1949] *A Modern English Grammar*, Volume VII. London: George Allen and Unwin.
- Johnson-Laird, P. N. [1983] *Mental Models*. Cambridge: Cambridge University Press.
- Kamp, H. [1981] A theory of truth and semantic representation. In Groenendijk, J. A. G., Janssen, T. M. V. and Stokhof, M. B. J. (eds.) *Formal Methods in the Study of Language*, Volume 136, pp277–322. Amsterdam: Mathematical Centre Tracts.
- Karttunen, L. [1971] Some observations on factivity. In *Papers in Linguistics*, Volume 4, pp55–69.
- Karttunen, L. [1974] Presuppositions and linguistic context. *Theoretical Linguistics*, 1, 181–194.
- Keenan, E. [1971] Two kinds of presupposition in natural language. In Fillmore, C. J. and Langendoen, D. T. (eds.) *Studies in Linguistic Semantics*, pp45–54. New York: Holt.
- Ladd, D. R. [1983] Even, Focus, and Normal Stress. *Journal of Semantics*, 2, 157–170.
- Lapolla, M. V. [1986] The Role of Inversion, Clefting and PP-Fronting in Relating Discourse Elements. Technical Report No. AI86-28, Artificial Intelligence Laboratory, University of Texas at Austin, July, 1986.
- Lewis, D. [1979] Score-keeping in a language game. In Bauerle, R., Egli, U. and Stechow, A. (eds.) *Semantics from Different Points of View*. Berlin: Springer-Verlag.
- Nakada, S. [1973] Pseudo-clefts, what are they? In Corum, C., Smith-Stark, T. C., and Weiser, A. (eds.) *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society*, Chicago, April 13–15, 1973, pp428–441.
- Pinkal, M. [1986] Definite Noun Phrases and the Semantics of Discourse. In

- Proceedings of the 11th International Conference on Computational Linguistics and the 24th Annual Meeting of the Association for Computational Linguistics*, Institut fuer Kommunikationsforschung und Phonetik, Bonn University, Bonn, 25–29 August, 1986, pp368–373.
- Pullum, G. [1989] Topic ... Comment: Formal Linguistics Meets the Boojum. *Natural Language and Linguistic Theory*, 7, 137–143.
- Prince, E. [1978] A comparison of WH-clefts and it-clefts in discourse. *Language*, 54, 883–906.
- Prince, E. [1981] A Taxonomy of Given-New Information. In Cole, P. (ed.) *Radical Pragmatics*. New York: Academic Press.
- Prince, E. [1986] On the syntactic marking of presupposed open propositions. In *Chicago Linguistics Society 22*, Chicago, Illinois, 1986, pp208–222.
- Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. [1985] *A Comprehensive Grammar of the English Language*. London: Longman.
- Reichgelt, H. [1986] Reference and Quantification in the Cognitive View of Language. PhD Thesis, School of Epistemics, University of Edinburgh.
- Reichman, R. [1981] Plain Speaking: A Theory and Grammar of Spontaneous Discourse. PhD Thesis. Published as BBN Report No. 4681.
- Reichman, R. [1985] *Getting Computers to Talk Like You and Me*. Cambridge, Mass.: MIT Press.
- Ross, J. R. [1967] Constraints on variables in syntax. PhD Thesis, MIT. Indiana University Linguistics Club.
- Sanford, A. J. and Garrod, S. C. [1981] *Understanding Written Language*. Chichester: John Wiley and Sons.
- Selkirk, E. O. [1980] Prosodic Categories in English Word Stress. *Linguistic Inquiry*, 11.
- Seuren, P. A. [1984] *Discourse semantics*. Oxford: Basil Blackwell.
- Sidner, C. L. [1979] Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse. Technical Report No. 537, MIT Artificial Intelligence Laboratory, June, 1979.
- Sornicola, R. [1988] It-Clefts and Wh-Clefts: Two Awkward Sentence Types. *Journal of Linguistics*, 24, 343–379.
- Stalnaker, R. C. [1974] Pragmatic presuppositions. In Munitz, M. K. and Unger, P. K. (eds.) *Semantics and Philosophy*, pp197–213. New York: New York University Press.
- Svartvik, J. and Quirk, R. [1980] *A Corpus of English Conversation*. Lund: LiberLaromedel Lund.

- Taglicht, J. [1984] *Message and Emphasis: On focus and scope in English*. London: Longman.
- Thompson, H. [1977] Strategy and Tactics in Language Production. In Beach, W. A., Fox, S. E. and Philosoph, S. (eds.) *Papers from the Thirteenth Regional Meeting of the Chicago Linguistics Society*, Chicago, April 14–16, 1977.
- van der Sandt, R. A. [1982] Kontekst en Prespropositie: Een studie van het projektieprobleem en de presuppositionele eigenschappen van de logische konnektieven. PhD Thesis, Nijmegen Institute of Semantics.
- van der Sandt, R. [1988] *Context and Presupposition*. London: Croom Helm.
- van der Sandt, R. A. [1989] Presupposition and discourse structure. In Bartsch, R., Benthem, J. and van Emde Boas, P. (eds.) *Semantics and Contextual Expression*. Dordrecht: Foris.
- Ward, G. [1985] The Semantics and Pragmatics of Preposing. PhD Thesis, Department of Linguistics, University of Pennsylvania.
- Werth, P. [1984] *Focus, Coherence, and Emphasis*. Croom Helm.
- Williams, E. [1983] Semantic vs. Syntactic Categories. *Linguistics and Philosophy*, 6, 423–446.
- Wilson, D. [1975] *Presuppositions and Non-Truth-Conditional Semantics*. New York: Academic Press.
- Yekovich, F., Walker, C. and Blackman, H. [1985] The Role of Presupposed and Focal Information in Integrating Sentences. *Journal of Verbal Learning and Verbal Behavior*, 18, 525–548.
- Zeevat, H., Klein, E. and Calder, J. [1987] Unification Categorical Grammar. Research Paper No. EUCCS/RP-21, Centre for Cognitive Science, University of Edinburgh, Edinburgh, December, 1987.

Appendix: The Corpus

The corpus consists of both spoken and written data. The spoken data is a combination of data taken from the Survey Of English Usage [Svartvik and Quirk 1980] and the author's own collection. The written data is extracted from the Lancaster-Oslo-Bergen (LOB) on-line corpus of British English texts. The total frequencies of examples are as follows:

Cleft Type	Written (LOB)	Spoken (SEU)	own	Total Spoken	Total
<i>it</i> -clefts	165	54	37	91	256
<i>wh</i> -clefts	72	63	36	99	171
reverse <i>wh</i> -clefts	38	106	16	122	160

1 Data from Written English

The LOB data are given below. Citations given with the examples indicate the location of the clefts themselves, not that of the surrounding context which is also included. The form of a citation is as follows:

C35 898

The first digit indicates the data file on the tape. Where the file reference C appears, for example, the appropriate file on the tape is c.dat. The material in each file is divided into numbered texts, indicated by the numerical field in the first group of characters (35 in this case). Each text has its own line numbers, which appear as the second group of digits. The line number in each citation given below is the one at which the cleft construction begins, although examples (which may, incidentally, be embedded within another construction) generally extend beyond the line at which they begin. In each case, the cleft sentence appears in italics.

It-Clefts (165 examples)

- A01 200 They would provide an ambulance service for the completely wretched — but it would not be too comfortable nor too easy to get. Answering jeers that *it was Labour which first put a ceiling on health spending and started charges*, Mr. Brown reminded the hostile Government benches that was done in 1950 because of the financial strain of the Korean war. In fact, the Tories made it worse now for the sick and needy than Labour had to make it in 1950.
- A02 104 If the threatened counter-revolution was not enough to bring the President back from his travels it might have been thought that the muster from the 13 States of the Commonwealth was an occasion worthy of his presence. After all *it was Mr. Nkrumah who suggested that this year the Economic Consultative Council should meet in Accra*. It has been left, however, to Mr. Goka, Ghana's Finance Minister, to do the honours as host, in which capacity he held a reception tonight in Accra's Ambassador Hotel.
- A02 220 Police leave has been cancelled and secret plans prepared to deal with the mass sit-down rally planned for Sunday in Parliament-square by the Committee of 100, the anti-nuclear arms group. *It was Mr. Butler who authorised action which ended yesterday in 32 members of the Committee of 100 being imprisoned for inciting a breach of the peace*. The committee's president 89-year-old Earl Russell and his 61-year-old wife were each jailed for a week.
- A08 48 Lewis converted with a fine kick. Oxford's best means of progress was by the boot and *it was in this manner that they secured their only success*. McPartlin and Stafford hacked the ball from halfway, Lewis fell and missed it and McPartlin went on to score, Willcox converting.
- A08 98 Having to concede weight forward proved too much for the seamen and they lost by two goals and two tries to a try. A lack of determined defence in midfield and casual defensive covering allowed Swansea to score tries, but *it was the greater experience and vigour of Swansea, with five internationals, which carried the day*. After their good display against Newport the Navy failed to reproduce the same form.
- A10 201 Arthur Willcox, spokesman for the firm which makes the plane told me : The plane has a built-in stereo tape-recorder which can play for the whole four hours it

will take to fly to Majorca. We are recording hits from the London shows on it. We understand the Duke likes them. I understand *it was Air Commodore Sir Edward Fielden, Captain of the Queen's Flight, who recommended the aircraft to the Duke after seeing it at the Paris Air Show.* This plane is purely a demonstration model, said Mr. Willcox.

A11 176 In his letter of resignation Mr. Singleton wrote to Mr. Murray : Rightly or wrongly I felt justified in correcting the impression in the statement issued to the press that the opinion of the council was unanimous *It is with sincere regret that I sever my association with the council on which I have always thought it a great honour to serve.*

A16 138 They are Guinness, Courage, Barclay and Symonds, Mitchells and Butler and the Scottish and Newcastle Breweries. And *it was Lord Boyd, vice-chairman of Guinness formerly Mr. Lennox Boyd who in his best front-bench manner yesterday launched a campaign at the Dorchester Hotel to make us drink Harp on a national scale.* Now nearly all the big brewers (Bass is a notable exception) are committed to the struggle for the new drinking market.

A18 172 Germany has done terrible things to the soul of man. Perhaps *it is sympathetic fear which prompts his friendliness to other people.* He is at home with scene-shifters, ASMs, and strangers who stop to ask about his high-powered sports car.

A19 56 Above all, they would never, never know that the New Wave, and it is the one thing that Mr. Coward can no more forgive than he can understand, is supremely successful, or that his own latest offering to Britain's ungrateful stage (Waiting in the Wings) is being withdrawn shortly, having failed, as they say in the profession, to attract an audience. Yet *it is Mr. Coward, too old nearly 40 years ago, mark you, who offers himself as the man to lead the poor, stumbling audiences out of the theatrical dark and into the bright, brave noonday where it is always perfect anyone-for-tennis weather, and where nothing as vulgar and squalid as a stove is ever mentioned, but where lots of nice, jolly, fun-giving adultery to the immense, brittle amusement of The Master is.* I think it is time that the case for the British theatre of today was made, and made loud and clear.

A20 209 The men are well trained and well deployed. But *it is the role of the whole N.A.T.O. army that worries me and our role in that.* While I am clear myself on

what that role is, I am not sure whether the politicians' statements are clear to the military generals and to Air Force chiefs.

A28 26 At the airport, Mr. Kennedy praised his host as a captain in the field in the defence of the West for over 20 years, adding that his leadership and sense of history were needed more than ever today. *It was Mrs. Kennedy who drew the crowds*, said police. The President stood bare-headed in his car to acknowledge the cheers, but Mrs. Kennedy, dressed in a pale blue coat and matching blue straw hat, was half-hidden from the crowds as she rode by in her enclosed car, waving and smiling.

A39 19 Anything from £30 to £80, he said. Oh, yes, for one performance. *It is therefore on real hard cash that he bases his argument that if Manchester is ready to pay the top price of, say, 25 s. a seat, and fill the theatre, Manchester can expect more of the world-famous ones always provided that her claims do not unhappily conflict with those of La Scala, New York's Metropolitan Opera, Vienna, Venice, Paris, San Francisco, and the rest of the world's leading opera houses.* The career of an international opera singer is not a particularly long one.

A40 145 But, despite the painstaking research which occupies the leisure hours of the keener enthusiasts, this anniversary nearly slipped by unnoticed. *It was 18-year-old Steyning Grammar School boy Michael Keeney, of Atherton, Jarvis-lane, Steyning, who came upon the fact that the Shoreham to Henfield railway, via Steyning, opened on July 1, 1861.* He got to work immediately.

B01 191 The same sort of good neighbour that he proved to be to the Jews fleeing from Hitler in the thirties. *It was Dr. Verwoerd who led a protest against admitting any of them because they would defile the national white stock.* It is impossible to make contact with Dr. Verwoerd in his nightmare world. It is this that makes illusory any hopes that he may be influenced to change course. The Archbishop of Capetown, Dr. Joost de Blank, has pleaded that South Africa should be allowed to stay in the Commonwealth.

B04 31 Not so long ago older folk were reminding young wives, harassed by shortages, of the good old days of abundance. Now *it is mother who picks up recipes from her daughter.* The dinner table is the best answer to the grumblers in Britain today!

B07 45 The objectives might be listed like this: that the system should be efficient; that it should be fair as between one taxpayer and another; that it should encourage personal saving and the wider spread of ownership of assets and property, and that it should contain the minimum disincentive to, and where possible should actively encourage, risk-taking, enterprise, exports and investment in efficient production methods. *It is the last of these four objectives about which we have heard most in the past year.* It is, therefore, with direct taxes on income (income tax and surtax) and capital (death duties and stamp duty) that we will begin.

B07 95 The second source of concern is the widely-held suspicion that a number of professional dealers in property and shares pay no taxes since their income is mostly in the form of untaxed capital gains. *It is from this suspicion that the main support for a capital gains tax comes.* The trouble with a capital gains tax is that it hits so many other things as well, including small savings and the smooth working of the capital market, besides being of low and uncertain yield.

B08 43 Then into the limelight stepped Selwyn. *It is not only Mr. Butler, the deserving candidate for Downing Street, who is in trouble.* So are many other prominent contenders for the Premiership in the radical sector of the party.

B08 59 And Viscount Hailsham, a radical Tory even if he would dislike being labelled a Left-winger, is down in the dumps of the whimsically named Ministry for Science. Cast a glance along the Right Wing: *it is there that success lies at the moment.* Lord Home is wielding immense power at the Foreign Office.

B10 167 Sir, I presume that Mr. G.H. Woolveridge's letter (August 23) is written in his official capacity, and *it is for this reason that I do not think it should be allowed to pass without comment.* Firstly, what does it cost a motor trader to assist in filling up an H.P. form and posting it?

B11 56 A junior clerk of 16 receives £230 per annum: does Mr. Kelf-Cohen expect him to work for him for £22 per annum, or a young man to return from the forces at the age of 20 and work for him for £162 per annum (£370—£208)? It should be remembered that until the implementation of the Guillebaud Report, under which railway rates of pay were based on the principle of comparability with those of comparable employees in other employments, railwaymen had worked for considerably debased rates of pay, and *it was they who had been providing the subsidy necessary for the running of the railways which are necessary to the*

economy of the country.

B16 17 During the summer of 1917 he dispersed a Communist rebellion with a whiff of grapeshot now described by such progressive historians as A.J.P. Taylor as a massacre. *It was a blundering general, with the active encouragement of the English and the French, who destroyed Russian democracy by attempting a right-wing putsch, which was suppressed without a shot but left the masses confused and distrustful of Kerensky.* This turn of events enabled Lenin to mount a counter-attack which the vast majority of Socialists tantamount to a majority of the nation resisted only with talk.

B17 45 Berlin could be an independent city and used as a home for the United Nations. It is true that, whatever happens, the Germans look like being left with a divided country, in itself a dangerous situation, but, as has been said many times before, *it is the Germans themselves who are at the root of all these problems and they must be satisfied with whatever terms their conquerors feel are necessary to maintain the peace of the world.*

B17 60 He submitted that in our society there was evidence that education was failing to keep up with the increasing tempo of materialism. *It is, of course, this aspect of the matter that is disturbing the Home Secretary.* One of the great tragedies of modern times is that our busy schools are kept at full stretch educating the young in the practical things which they will need to make their way in a highly competitive society.

B17 75 It was a heavy, distasteful task that fell to Mr. Frank Foulkes yesterday. For Mr. Foulkes is, of course, the president of the Electrical Trades Union, and it was in that Union, and *it was in that capacity that he announced the results of the elections for the membership of the union's general executive, in which the Communists have suffered an overwhelming defeat.* Naturally Mr. Foulkes, who is himself a Communist, put as good a face on it as was possible in the circumstances, but it did not amount to much.

B19 5 Lord Amory is to head the Central Advisory Council for Education during its consideration of the 12 to 16 age group in our schools and further education institutes. *It is within this age group that outlooks are formed and decisions are taken that lead to lamentable waste of young people who could make a valuable contribution to our national life and who do not, for the most part, make the best*

of their own lives. Lord Amory's long-standing interest in youth particularly in the young teenagers now to be considered will be of great value to the Council as will his personal experience in medium-sized industry in which large numbers of youngsters must find their first jobs.

B22 48 Mr. de Freitas became M.P. for Lincoln in February 1950 and his term will have extended to 13 years and eight months by October 1963. To return to Sir Walter Liddall: *it was in July 1944 that he set up his Parliamentary record by beating the term of office of Mr. Charles Roberts, who was Liberal M.P. for Lincoln from 1906 to 1918.* But the all-time record is one of 20 years, held by Colonel Charles Sibthorp, an early Victorian Member, and one of a number of gentlemen of that family who at one time or another represented Lincoln in the Commons.

B26 164 The Labour Party is fundamentally more united than before. *It is not the withdrawal of the Whip that causes new disunity :but the deliberate defiance by five members of decisions by the party.*

C01 91 It struck me that Mr. Bennett's ideas in this piece were not so much succinct as slender. Perhaps *it was for this reason that the work seemed somewhat pale in character,* a criticism that certainly cannot be made of Berg's very rarely heard Three Orchestral Pieces, Op. 6 each bar of which, even the most derivative, is impregnated with the composer's personality. The cruel acoustics of the hall played havoc with textures which are unusually hectic and congested, but Mr. Del Mar's heroic labours conveyed a clear impression of the succession of catastrophes which seems to be the work's natural mode of expression.

C01 190 Byron's marriage, the reasons (real enough though embroidered later) for Lady Byron's leaving him, the scandal of his love affair with his half-sister, Augusta Leigh, the question of the paternity of Medora Leigh her daughter, the long inquisitorial persecution of Augusta by Lady Byron (who seems to have been as neurotic as the most ghoulish novelist could wish), the patient ferreting for evidence to add homosexuality to incest as an extra nail in his coffin, the unspeakable treacheries of Lady Caroline Lamb, the scarcely less heinous treacheries of Augusta—it is *the Lovelace Papers, surely, that deserve to be called a 'pest bag', not Byron's consumed Memoirs, which at least would have possessed the merit of being well and entertainingly written.* Equally important have been the Hobhouse

Journals, a vast mass of material partly in the British Museum, partly in the possession of the Hobhouse family in Somerset.

C03 37 In Fanny the pregnant girl is befriended by an old man. Here *it is a young homosexual, estranged from women but yet moved by a strong maternal instinct to the unborn child as much as to the expectant mother, who acts as a protector and comforter to her in her hour of need.*

C06 142 Result: Mr. Hudson and lady love Lollo find themselves playing chaperon (Brenda de Banzie, the official one, has broken a leg) to the girls, who have just been joined by a Jeep-load of boys. *It's hereabouts that the budgie takes to the bottle,* but I don't think it was through boredom. The film is funny enough in places and has a line or two of painful home truths thrown in.

C07 46 Lautrec's vision of his women is, I think, the outcome of some such ambivalence as this: on the one hand, celebration of their easy animal vigour and grace; on the other, celebration of the knowledge that they too would fall into decrepitude. *For it is not a present state of decay that Lautrec presents as a rule, but only an intimation of decay.* He isn't at all Swiftian about women: he doesn't, getting close, rejoice in recoiling from their enlarged pores.

C07 88 Rodin's ghost will not be laid. *It is that old master's energy and rugged form, rather than his aspirations, which have influenced two of the three conspicuous sculptors this week:* Ralph Brown (Leicester Galleries) and the American Jack Zaja (Roland, Browse's). Ralph Brown began as a social realist sculptor infusing tenderness into a gawky mother fondling a child, an infant bowling a hoop.

C10 151 Radha and Vasanti are graceful, too. *It is how their brother Mr. Kumar got on stage that beats me—unless, of course, he is really Peter Sellers.* From his performance, I guessed that, watching his kid sisters perfecting themselves in their art, he suddenly couldn't bear not to be in on it too, and finally forbade them to appear without him.

C12 68 Of other Greek cities only Athens and Constantinople have equally powerful associations, and the worlds of Alexandria and Constantinople are, of course, utterly different from the world of fifth-century Athens. *It was out of the world of the Greek dispersal that Cavafy created his personal mythology—a world both of triumph and disaster, a world of courage, of humour and of irony.* Cavafy was the

first modern Greek poet who contrived to be patriotic without being romantic, and his method was to stand at 'a slight angle' to what is assumed to be the universe of history.

C12 164 He reminds his French readers of Rimbaud's obscene parting rites in the home of an acquaintance and explains that Lawrence's own ungrateful mocking of those who had helped him was only to be expected in a great artist. Someone as courageous as Lawrence in following the promptings of his own intuition is bound to inspire the jealousy or the envy of those who are more timorous and conventional and *it is probably for this reason that so few of his critics, whether or not they have known him personally, have been capable of a truly disinterested assessment of his character and genius*. M. Temple's short study of the life and works is on the whole eulogistic and he defends Lawrence vigorously against some of the charges that have been brought against him in the past: ...

C12 174 ...that he was a precursor of Nazism, that he sentimentalized the noble Mexican savage, that he suffered from the neuroses described in Murry's *Son of Woman* and that he earned money to which he was not entitled by publishing Maurice Magnus's *Memoirs*. *It is only occasionally that he gives the impression of not wanting to sound too impressed, as, for example, when he mentions in passing the numerous (unspecified) puerilities in Lawrence's daily life and in many of his books*. M. Temple makes good use of the available biographical information.

C12 181 He also quotes lengthily and well from Lawrence's letters. If one is forced to conclude that he seriously misrepresents both the life and the work of Lawrence *it is not therefore because he is swayed by any deep prejudice or because of any particular inaccuracy (his worst inaccuracy is to describe Ursula in *The Rainbow* as Tom Brangwen's daughter)*. The principal defect of this book is that it is written in a style which will convey to the reader little or nothing of the resemblances between Lawrence's inner life and his own:

C13 66 Criticism has been misguided and has underestimated the book because it has 'seized upon the superficial content of the novel and confused it with the story it is really telling'. But *it was precisely because the 'primal mythic adventure' could not form the total substance of a novel that Lawrence was driven to invent the paraphernalia of a political and religious movement led by Ramon which Mr. Kessler rightly regards as superficial*. It is impossible to 'rescue' the myth from

the novel.

C15 63 The action covers one day in the lives of the occupants of a pleasant country house who find they have a body on their hands shortly before the arrival of an important foreign diplomat. At all costs this must be covered up so that the important conference with the V.I.P. can take place, and *it is in this endeavour that the plot develops, drawing into it a number of mysterious suspects*. Introducing 10 years old Wendy Turner to the screen as the daughter of the household, the film also enables David Nixon to make a guest appearance.

C17 164 That is why John Cassavetes came to England to find someone who would take a risk on something new. *It was the directors of newly-constructed British Lion, who have got faith in fresh faces, talent, ideas and letting people try them out, who saw 'Shadows' one evening and immediately offered Cassavetes the money for world distribution rights*. I feel sure they won't regret it, from the prestige or financial angles.

E01 96 During the sixteenth century a considerable amount of crochet was produced in the convents of Europe. Without doubt *it was the nuns who carried the craft to Ireland*. There it was developed into quite an elaborate and distinctive form with rosettes, leaves and lace fillings.

E01 104 One can remember with horror the pictures of the overcrowded and over-embellished drawing-rooms complete with heavy crochet antimacassars, mantelpiece covers with a fringing of clumsy bobbles and numerous other crocheted pieces. *It was some years after the Victorian period that designers realised the potentialities of the craft, and crochet was revived with designs suitable for contemporary trends*. Crochet today has a variety of uses, and has even stepped into the field of high fashion.

E04 87 If you do not have a wooden floor on which to anchor the moulds, make a frame of rough lumber as shown in Fig. 2. The main idea is to have the moulds standing as rigid as possible, for *it is on these you will be building and shaping your little craft, upside-down*. For cheapness I recommend using Douglas Fir Plywood from British Columbia.

E07 45 After extricating itself from the membrane, the young mantis has to dry itself. *It is at this stage that the helpless insect is in great danger—especially from ants*.

I found that neither sticky bands nor tanglefoot deterred Kentish ants.

E07 114 Despite the steady influx of substitutes since the end of the war, clay pots are still in greatest demand. *It is mostly in the smaller sizes that the substitutes have made headway, and although numerous small clay pots are used each year, they are fewer than they were 10 years ago.* In the larger sizes there is still no real substitute for clay that can be obtained at a comparable price, except the bituminized paper pots which are used on commercial nurseries for growing tomatoes and chrysanthemums.

EO9 75 Miniature cedar trees are used to block out the original value. *It was not until 1947 that the Nahr el-Kelb scene again appeared on stamps.* In that year four airmail stamps in offset lithography were printed in Beyrouth to commemorate the evacuation of all foreign forces from the Republic.

E12 172 Much credit for this effort goes to the Secretary, Mr. Aubrey Pugh, who carried the bandsmen and instruments in relays in his own van to save the band's fund. *It is due to his financial and untiring help that the band is continuing to function as they are not supported by any industrial firm.* Deiniolen and District (J.E. Williams) played carols in their district during the week before Christmas.

E13 11 Recent articles of 'Steps to Success', have been for beginners in the Four Standard English dances, and we have covered efficient to enable the non-dancer of a few months back to now dance around in the average ballroom without feeling self-conscious and with just enough figures to avoid the monotony of repetition. In the dancing school *it is usually when this stage is reached that the pupil is asked to move on to a different class of instruction; either to an intermediate class, where an added variation is taught at each session, or to a Bronze and Silver Medallist Class, where time will be mainly spent in improving the style of dancing and the execution of the figures already learnt.*

E15 44 But apart from these domestic problems the question of the prolongation of the season depends on agriculture. Foxhunting, after all, is a 'trespass by courtesy' and since the courtesy is on the part of the farmer, *it is the latter's interests that finally decide the matter.* Those counties that have a bit of hill or downland are often able to continue operations after the vale is closed.

- E23 147 Another Spring flower, the iris, is sometimes called 'The poor man's orchid'. *It is not the colour or the texture of iris petals that fascinate me, but the fine detail of their exquisite shape.* I love them all, the Dutch and English varieties, and later on the Flag and Bearded types among which there are such exciting colour combinations.
- E24 108 'Written composition,' it said, 'is generally begun too soon and practised too often.' *It is only tradition and obsession that demands a weekly piece of writing from each child in the class.* If we look ahead a little to the work of the majority of our leavers, we must admit that few will need (nor will they wish perhaps) to have to write a formal letter.
- E26 17 Then, after you have ascertained whether it is to be tea or coffee and made sure of any other details, you can go to bed yourself, to sleep with an easy conscience. *It is not everybody that likes to be given breakfast in bed; on the other hand, because it is so many people's idea of bliss, I would like to say something more about it.* The bliss can be considerably marred if the tray is overcrowded or ill-arranged so that in order to pour out it becomes necessary to move things, to play a sort of game of chess in bed.
- E26 87 A nice old-fashioned housemaid, labelled by cap and apron, is easy enough; when you leave you will give her your little present as you thank her for looking after you. *It is the 'lady who obliges' that can confound you; on that point, the simplest way is to quietly consult your hostess.* In the old days a young woman was not expected to tip men servants; nowadays if a chauffeur meets you and takes you to the station, you will want to show appreciation.
- E26 110 Young people of the present generation have conditioned themselves to what is sometimes called background noise, and can carry on conversations, read and play games against a radio programme—even against two from different stations, without apparent inconvenience. *Maybe it was the early training which expected a complete cessation of noise of any kind when music was being performed, that causes me and others like me to find it quite impossible to talk or listen when there is 'background noise'.* To be a little considerate about radios and gramophones and noise generally is rated highly among good manners.
- E28 19 'Rental Values'. *It was not until 1836 that the basis of assessment was defined by statute.* The Parochial Assessments Act of that year stated that all rates were

to be based on the rent at which rateable hereditaments might reasonably be expected to let from year to year.

E29 9 To mark the event a dinner dance was held at the Scotch Corner Hotel and upwards of 200 farmers, friends, and staff were present. While *it was a second-hand machinery sale that was celebrated*, the main feature in the ballroom was a new Massey-Ferguson tractor, which shone quite brilliantly under the spotlight. We understand that, while it excited much attention, it did not intrude in any way on the dancing.

E30 72 Doubling the selling space to 700 sq. ft. by adding a new section of over 300 sq. ft. was not to be the greatest expense. *It was the new fixtures and fittings to fill this space that would be costly.* But Roy Beddis solved this problem by building all the new sectional shelving, eight dump bins, and the large gondola from materials he bought for a total of 50.

E35 51 It is advisable to keep in reserve a small amount for advertising during March / April to catch those people who do not decide where or what they are going to do until later in the year, or possibly are wanting to see what the Budget is going to do for them. *It is at this time when the public are thinking of planning their forthcoming annual holiday.* It would be extremely wasteful to spend money on promoting summer holidays in the middle of June or July.

E35 124 However, nothing materialises, no positive action takes place, no orders are negotiated, until the prospective purchaser is in contact with a travel agency or a transport organisation. 'People taking charge'. *It is now that people become involved ...*

E35 124 ...and *it is people, with their idiosyncratic ways, patterns of behaviour and thought processes, who make or mar a sale.* The human personality takes effect.

E36 130 While this general picture of the way the farm was run will be of interest to the practical farmer, *it is the economic aspect which the experiment was undertaken to test.* This aspect is treated in detail in the next section of this report, but in considering it three points should be borne in mind.

E37 92 This illustrates how wrong it is to attach too much importance to the price of a ration. *It is the feeding value which counts.* A food costing £30 a ton may be dearer to buy than one at £28 a ton but it can be cheaper to feed.

- F01 9 Both the item and the sketch can then go into the store. In the mind *it is the stored items which constitute our memory ...*
- F01 10 ...and *it is the stored sketches which constitute our understanding*. Thus reading for understanding means taking items of knowledge to pieces as we read them and seeing how the pieces are connected.
- F01 143 To begin with it is a pattern of brain-processes, just as physical as the water-drops. But it is experienced as a cloud, as a whale, as a camel and so on. We cannot dismiss these as 'illusions' for *it is just the occurrence of such illusions that we seek to explain*—besides why is it illusory to see the thing as a whale but not illusory to see it as a cloud? And how did Hamlet know it was 'really' a cloud?
- FO5 77 Then, as the United National Independence Party, Mr Kaunda's group promised independence by October 1960, which was rash. The potentialities for conflict existed in all three countries, but *it was in Nyasaland that the nationalist organization developed its greatest energy*. The Nyasaland Congress had been formed in 1950; the institution of Federation three years later provided it with its *raison d'être*, and in 1958 it received at last the genuine leadership and stimulation it had awaited.
- F07 39 The peasant in those days, as ever, ate sparingly, but generously enough in his own fashion, save at feast times, when he, too, let himself go; and *it was from his simpler food that the later renaissance of cooking was to come*. Epicures and gourmands, sated by the unending procession of dishes from those mammoth kitchens of the 18th century—that amazing epoch of grossness and delicacy—sought inspiration at last from the dishes of the country, and, instead of gorging the eye with magnitude, began to understand the value of intelligent selection and comparative simplicity, though nowadays their simplified meals would seem quite monstrous.
- F10 137 So delicate is the balance that it can be tipped by a slip of the tongue or by some development that shakes the confidence of the timid or uncertain. *It is on this uncertain group that the choice of government rests*. It's a grave thought.
- F10 199 Film-strip talks play an important role in many schools all over the country, and they should play a much more important role in the N.C.L.C. than they do, because people learn through their eyes as well as through their ears. Besides,

if one looks at the papers that have the biggest working-class circulations it is *they which have the most pictures*. One of the latest history film strips available is 'George and the Revolutionary Wars (1760-1815)', published by Common Ground, Ltd.

F16 188 Nor will they eliminate the mounting frustration, boredom loneliness and tension, felt by an increasing number of people. Perhaps *it is the knowledge of this fundamental truth—that real happiness and satisfaction is found in doing for others, that enables councillors to labour on year in and year out, unpaid, unrecognised, in what must appear to others to be a thankless and unrewarding task*. Does this sound priggish, evangelistic, dull?

F16 212 From his earliest years he was an active trade unionist and Labour Party worker. However, *it was not until 1920 that I first knew of him*. He was in Leeds and I in London. In that year his Union—the Tramwaymen's Union—amalgamated with others to form the United Vehicle Workers' Union (U.V.W.), and my Union, the London Carmen's Trade Union amalgamated with others to form the National Union of Vehicle Workers (N.U.V.W.).

F19 99 Prayer, psychic abandonment and the many kinds of devotional exercises induced in primitive man, accompanied, as it did in all his descendants, the condition known as spiritual elevation and exhaltation, followed by the more advanced stages of inspiration and ecstasy. *It was only in these later spiritual phases, that the human mind was able to step across the threshold of material thought into the sphere of the immaterial or supernatural world*. In these phases only, could man leave his objective consciousness entirely behind him.

F20 57 The farm which was the Company's first purchase, High Wallabarrow, lies on the Cumberland bank of the Duddon, opposite to the hamlet of Seathwaite in Dunnerdale, and adjoins the National Trust property at Wallabarrow Crag. The farm and its fell land are within the area which the Forestry Commission had declined to preserve from commercial afforestation and *it was to anticipate a purchase by the Forestry Commission that this farm was acquired by the Company*. The landscape is a fine example of the beauty characteristic of the Duddon Valley, and the farm house, though in itself a small one, is delightfully situated.

F21 144 On his death at eighty-four he bequeathed his notes, comparative charts and unreturned library books to his son, then aged fifty-six, with the request that

he complete the task by knocking the book together. The son occupied twenty-two years very pleasantly in reading through, revising and annotating his father's notes, and *it was a grandson, a very well-rounded personality of forty-eight with no leisure-time problems, from whom I heard the story.* Here is wisdom indeed, when a man can cater not only for his own middle-age and old age relaxations but for those of his descendants as well.

F25 13 The rank and file of the expedition were drawn from many parts of the Emperor's wide domains, and even from lands beyond. A particularly large and well-armed contingent came from Portugal, and *it is to one of these Portuguese adventurers, known as the Gentleman of Elvas, that we owe the most circumstantial first-hand account of the expedition.* Amongst volunteers of other nationality we find mention of a French priest from Paris, Biscayan carpenters, a Genoa master-craftsman who could construct anything from a bridge to a brigantine, a Spaniard reared in England, and even an unnamed Englishman whose skill with the long-bow matched that of the Indians.

F25 43 Florida seemed a promised land indeed. *It was not until some days later that the first Indians were encountered.* Amongst them was a man, all but indistinguishable from the natives, whom the Spaniards almost rode down.

F27 169 How the outlaw was rewarded is told in the Tale of Gamelyn: the knights of the county might conspire to cheat him, but his villeins were faithful even in the hour of extreme misfortune:

Tho were his bonde-men sory and nothing glad
When Gamelyn her lord
wolves heed was cryed and maad.

It was to protect them against the oppressions of their new master that Gamelyn came to the Moot Hall, where he was arrested and bound by the sheriff. Whether he is like Gamelyn a knight or like Robin Hood a yeoman, the outlaw hero of the fourteenth- and fifteenth-century stories is the friend of the poor: he is not consistently the friend of the knight.

F29 174 The last reigning monarch to attend the Assembly was actually James, before he became the King of 'Great Brittain,' and before the appearance of his Authorised Version of the Bible; and he did so in order to discipline the members, not to praise or encourage them. *It was he, too, who instituted the office of High*

Commissioner, so that the Crown could keep a good eye on the proceedings; and ever since Jacobean times the Sovereign has been represented at the Assembly by a royally appointed representative. The office of Lord High Commissioner is now more ornamental than functional, at least in the sense that the holder is no longer a 'spy' in the pay of the Crown, which itself has changed beyond all recognition and is completely above politics or religious faction.

F31 123 Just as there is no specific cure for eczema, so there is no one specific cause. Nevertheless, it is regarded as an allergic reaction, although *it's only in rare cases that a particular substance can be detected to which the child is allergic.* Some believe that eczema is caused by emotional factors, even in the youngest baby.

F33 102 He'd rushed to the surgery and was breathing heavily. But *it wasn't that which disturbed me. ...*

F33 103 ...*It was the time he took to recover.* If you're under 30 you should get your breath back in one minute.

F34 43 The Post-Impressionists felt that this impartiality was itself a limiting thing. They held that *it was the painter's feelings about a scene that should be expressed, not just the light that reflected from the scene.* With this in view they permitted themselves to exaggerate any quality which they found exciting—they claimed the right to distort the facts according to their own feelings.

F35 36 Ruin and destruction seemed inevitable. *It was then that a beautiful young girl, named Lore, accompanied by a crowd of small children, offered to go out to meet the Colonel and to beg pity for the town.* But before the plan could be realised the Swedish troops had entered the city, ready to destroy it.

F37 13 22) Years later, after I had come across other versions of the song, I discovered why. Mr. Crouch, the blacksmith, as a child, had been in the party that gave the song to Sir Ernest Clarke at Hinwick Hall in the first decade of the century, and *it was Sir Ernest who had sent it to Lucy Broadwood. ...*

F37 15 ...*It was the Church family of Biddenham who first brought home to me the fact that there were other versions of this carol still known, and sometimes still sung, in Bedfordshire.*

F40 41 The distant motor-cycle caused me to give a momentary reflection on the calamity of road accidents. The barking dog made me pause to find out if *it*

was one of my own dogs barking, and if so for what reason. By contrast with our continual alertness to noises and their meaning it is possible at times so to lose oneself in preoccupation as to be oblivious to outside sounds.

F44 82 Russ Conway's injured hand has given him time to think—and the result may well be that a new field of entertainment will open up for him in 1961. *It was towards the end of November you may recall, that Russ had to withdraw from the London Palladium revue 'Stars In Your Eyes'. A fall in which he had suffered a severely bruised hand and wrist was the cause.*

G03 110 Alf Barton, who in was presented with a book on the life of Marx, and a gold purse in recognition of his work for the movement, was later known to say that it was not necessary to understand Marxism in order to understand socialism, though at this time he was a keen member of the B.S.P.. George, however, never deviated from his belief that *it was the economic basis of society which needed to be changed*, for the conditions of the people were appalling, there being only slight alleviations.

G04 28 The work was growing pretty heavy, and we managed to get a stipendiary layman who could help among the children and young people. *It was while I was at Tatsfield that I first visited Oberammergau in Bavaria to witness the Passion Play.* The place and its people were to play an important part in my life.

G05 39 This time, mercifully, there had been little destruction but warlike atmosphere was not entirely lacking since, through the town, ran one of the 'Red Ball Highways', those one-way highspeed supply routes along which by day and night thundered the endless convoys of giant American supply trucks carrying supplies from Normandy to the battle-fields. *It was seemingly not only humans that derived comfort from the roar of engines, for it seemed to have positively intoxicating effect upon the nightingales that appear to exist in Rheims in great profusion.*

G07 38 The papers were reopened in the presence of leading Missolonghi officials in order to make sure that no recent will was amongst them. *It may have been then that Trelawny contrived to do his copying.* Considering that Pietro was not above twenty-three years of age when he undertook a load of heavy responsibilities, his conduct reveals him as one of the most intelligent as well as the most sympathetic of Byron's entourage in Greece.

G08 166 ...the Viceroy was Lord Hardinge who 'lacked Lord Minto's enterprise, and was in every way a more conventional and less imaginative man', while the Secretary of State was Lord Crewe, much absent from the India Office on account of ill health and other duties in the House of Lords. Ritchie was permanent head of the India Office during most of this time, and it is not surprising that Sir Mackenzie Chalmers considered that *it was only through Ritchie's great ability and devotion that the Government of India was enabled to pull through the serious difficulties of those years; that Sir Henry Dobbs wrote that Ritchie had very great influence on affairs in India and saved the Government from many mistakes; that Sir J.R. Dunlop Smith considered Ritchie's death a blow to India not easy to measure ...*

G12 18 Does 'Naval Officer' want our fleets to lie alongside the home ports, Gibraltar or Malta, for nine months in the year? *It is not every naval officer who is afraid of battle exercises, or manning and arming ship, or of sea trips between nice places.* If 'Naval Officer' chooses to present one side of the case to the British public, surely the views of the majority may have a hearing also.

G12 25 *It was about now that I took action against Their Lordships themselves in the matter of the yearly Examination in French of Junior Officers Afloat.* My diary simply records: French exam. Had hoped to do well but they asked what were the pronouns which correspond to the adjectives 'ce, cette, ces, son, nos, leurs.'

G13 99 Mr. Wells considered himself to be very nearly an average man. If he was at all abnormal, he supposed that *it was 'only by reason of a certain mental rapidity.'* Be this as it may, the outbreak of hostilities evoked much the same response in Mr. Wells as in many other Englishmen.

G14 53 The confidence tricksters, it seems, consider it axiomatic that no wholly honest man can be regarded as a likely victim of the confidence trick. *It is not the mere fools that the confidence men successfully delude.* It is, in their pregnant phrase, the 'larceny in the blood' of the victim which results in his victimization. And that was how Hitler operated—exploiting and using as his leverage the 'larceny in the blood' of innumerable politicians in every country who wanted to believe that here was a man who really had found a way of making diamonds out of plastics; a way, that is to say, of making a quick profit out of an illicit sale of the Western soul.

G15 129 The massacre of the Jews in his beloved Ukraine, and the uncertainty as to what might be the fate of the yishuv, intensified his unhappiness; and his malaise adversely affected his physical health. Paradoxically, *it was during this period of acute distress that he made for the first time a direct contribution to the shaping of the policy of the Zionist Organisation.* Thanks to his intimacy with Dr. Weizmann, he was kept informed from the outset of the steps which were taken during the war to win the sympathy of the British Government and British public opinion for Zionism.

G15 182 Now, at the age of 70, he felt that he had earned the right to retire on a pension which would enable him to live in reasonable comfort in the land of his dreams. For unknown reasons, over a year elapsed before the necessary arrangements could be made; and *it was not till the end of that year that he was able to leave London for Palestine, accompanied by his wife and their son and daughter-in-law.* He preferred to live in Tel-Aviv, which was a creation of the new spirit of Jewish nationalism, rather than in the Holy City of Jerusalem, to which the aura of medievalism still clung; and the Tel-Aviv Municipality built him a house next to the Gymnasia Herzlia, the first all-Hebrew secondary school of modern times.

G18 15 At times the Rector was justifiably concerned at Wesley's indiscreet religious zeal, but realized his merits, and in June, made a donation to the work of the Castle, a gesture by which Wesley was obviously touched. Wesley had been recalled to act as tutor to the undergraduates, and *it was as a teacher and preceptor that he had returned into residence in November.* He was already well-read in the classics and in divinity.

G18 43 Green's father lived at Shipton, where Wesley often took the service for his friend, the former Lincoln undergraduate, Joseph Goodwin. *It was probably through Wesley's efforts that Green came to Lincoln.* He was soon calling on Wesley, who lived in rooms just above him in College, at ten every morning, presumably for tuition.

G20 2 The days grew shorter, but given fine weather, another crop could still be gathered. *It was on 21st October, as King of the Belgians, that he made a confession to the Archduke John: 'The Prince of Prussia has also written to tell me that you regret I have tied myself to Belgium. I too sometimes regret that my part in the East was taken from me.'*

- G20 38 There is no difficulty which does not alarm him, no obstacle which does not stop him, no gesture which does not prove that he would have brought to Greece disgust, pusillanimity, and the perpetual regret of having abandoned his so-called chances of the eminent position of Regent of England. *It is this Regency that he will never obtain, above all now that he has crowned his shame like this ...* Such a sovereign would have done damage to royalty.' And this scorn and anger were echoed by the correspondent, quoted in the Memoirs of Baron Stockmar ...
- G21 193 It was in fact a kind of literary 'class distinction', a superior quality which only the select were capable of appreciating. *It was not the matter presented by the author so much as the manner that counted.* The reader learnt to be sensitive to the shape of a sentence, to the use of 'master words' round which an author like Stevenson would build significant paragraphs; and to admire those splashes of colour that were almost purple.
- G22 60 I remember how I would open it and read the first words: 'You know, my dear little Arthur' and then turn to the last page and read the last words: 'I hope it will help you to understand bigger and better histories bye and bye.' I don't know if it was 'Little Arthur', but most certainly it was little Miss Gray who helped me to that understanding, awaking in me, sublimely unconscious, interest and energy for tackling these 'bigger and better histories' in later years. One of our lessons was to read aloud.
- G22 139 My blacksmith too, had 'large and sinewy hands'—'swiney' as one of my own children misread it—and often did I stand and watch him shoeing a horse, and was allowed to put my small hands on the bellows and help blow the fire. So *it is of my Clovenfords blacksmith, dark-eyed and black-bearded, in his smithy among the hills, that Longfellow brings back the memory.* At ten o'clock Miss de Dreux rang the big brass bell in the hall.
- G22 176 Before very long we were reading *Un Philosophe sous les Toits*—I cannot remember the author, but I know I had a sort of affection for that old philosopher and his meditations under his roofs. *It was dear Miss Bogen who gave us our first German lessons, only vocabulary, no books.* She was a sweet, kind creature and we all loved her.
- G24 121 But it was a sad parting, and I always missed them through the years of aero-engine and car designing. *It was, in fact, while I was working on locomotives*

at Doncaster that I became a motor-bicycling enthusiast; and I certainly got more pure fun out of the motor bicycle than I ever got from any of my cars, although I willingly accept that sport on two wheels is essentially for the young, and for me it was only a sport, with no commercial purpose behind it.

G25 110 There was the garrulous Bracey with the red face, monotonous voice, and stiff knee covered with wounds, who sat on the bed and told his story: he said that every anaesthetic took six months off a man's life; he had already had sixteen, so that meant he had lost eight years—and there were still more operations to come; yet that was better than being like Cain or Thompson who had each had a leg off, or better still than the little Canadian whom I often carried about in my arms because he had lost both his legs. But *it was Max the tall Irish Guardsman with his thin waxen face and black hair who distressed me more than any of the others, as he stooped and coughed as he walked about.* He had a huge wound in his chest which the sisters washed out with long tubes and hissing fluid, and then he coughed and spat as he tried to get his breath.

G26 14 But I was too busily engaged on the process of rehabilitation to want her company, and she was a woman—suspect as such, and further suspect owing to her happy association with holy writ that linked her with my father. *It was not till the middle of the week that I began to welcome her, caring for her until Saturday night.* Then, with the sound of the first church bell on Sunday morning, all women were suspect again; and as the hour in the box-pew remorselessly approached—the hour of avoiding looking at Milly, at the same time trying to reconcile her with my visual world—I knew it would only lead to the hour of afternoon when the sunlight froze on the tops of the trees, immobilized as I by the bible.

G27 146 And his 'pair of petticoats' for public inspection, though there might be another petticoat in the emotional background, were now Georgina and Mamie—and what could be more outwardly respectable? *It was they who went to the great farewell dinner held in London when he was invited to visit America for the second time.* His visit was a tremendous success, and it was they who welcomed him back to Gad's Hill upon his return. Georgina was not in the company of Dickens when he met with his first railway accident at Staplehurst, as were Ellen Ternan and her mother.

- G28 144 Never did he refer to this letter to the Halle Society, demanding my excommunication, at any of my subsequent meetings with him, not even during our day by day, night by night expressions of brotherly love in Australia. *It was round about then that he told me he was about to form a new orchestra in London.* 'But where,' I asked, 'where do you hope to find the players—the B.B.C. Orchestra has taken the best?'
- G29 64 The tremendous debt which Helen and blind people the world round owe to Anne Sullivan is beyond computation. For *it was Anne who rescued Helen from her world of darkness and misery, and enabled her to bring deliverance to countless fellow sufferers.* Anne was born in poverty, and her eyes were infected from birth.
- G29 69 Her mother died when Anne was eight years old, leaving three children who were placed in the workhouse. *It was here that Anne spent the next four years of her life, being allowed no social contacts save that of fellow paupers.* One of them told her that blindness entitled her to go to a special school, but no one was interested in the education of a blind pauper child until Anne literally threw herself at the feet of the chairman of the visiting committee and pleaded 'I want to go to school.'
- G29 95 In complete self-effacement, sweeping all self-pity aside, she gave herself to Helen, working tirelessly to open lines of communications between the imprisoned child and the world of people and nature about her. *It was the day after Anne Sullivan's arrival that Helen learned the finger language for the word 'doll'.* Anne spelt it into her hand very slowly and deliberately, and got Helen to imitate.
- 31 157 So Mr. Baldwin took the front of the stage, which he was to share with the King, others, in the background, till the play was done, for, as His Majesty phrased it, they were to settle the matter alone. *It is the King who serves as ceremonial figure-head for his country.* It falls to his Prime Minister to speak on behalf of England.
- G32 42 'I was filled with questions', he says, 'and I had to carry these questions about with me unanswered. *It was thus that I reached my eighth year.*' During this year the family moved to Neudorff in Hungary, and here they remained until Rudolf Steiner was seventeen.

- G32 52 With them, he took part every year in the vintage and with their children he went to the village school. *It was through the assistant master at this school that the first great event of his life took place—an event that, he believed, influenced the whole course of his development and of his future work;* it was the discovery, in his teacher's room, of a text book on geometry. He was allowed to borrow it and through it he felt the deepest satisfaction he had yet known, for by this science he found justification for his own assumption that the reality of the unseen world is as certain a fact as the reality of the physical world.
- G36 90 The ability of a literary mode to expand into others is often taken as a sign of vitality, and it is true that between them fantasy and science fiction have gobbled up most of what was left of the horror story without much injury, but I cannot feel that the injection of these thriller ingredients is likely to lead to much beyond blurring and dilution. *It is not by capturing more territory that science fiction will improve itself, but by consolidating what it already has.* Such internal reconstruction would do well to start with an attempt to bring sexual matters into better focus.
- G38 47 You can, if you wish, class all science-fiction together; but it is about as perceptive as classin the works of Ballantyne, Conrad and W.W. Jacobs together as 'the sea-story' and then criticising that. *But it is when we come to the second distinction, that made among the sheep or within the pale, that my system would differ most sharply from the established one.* For the established system, the difference between distinctions within the pale and that primary distinction which draws the pale itself, can only be one of degree.
- G40 21 The Indians were too few in the land to introduce Prakrit or any Dravidian tongue as the language of conversation, but the court Brahmins brought religion and learning and furnished the primitive Malay with his first abstract terms, terms still used by the Muslim Malay to denote religion, fasting, heaven, sin, life, language, time, name, prince, property, thing, a fine, work and so on. *It is this background that gave the Malay stories from the Jataka tales, Bidpai's fables and the Katha Sarit Sagara or Ocean of story, carried down the centuries per ora virum, until they were written down and published in modern times.* Most of these stories are known throughout South East Asia and there is Buddhist influence in folktales.

- G40 41 By then Malays were instructed by a famous theologian writer in Malay that the Ramayana might be condemned to the rubbish heap provided the name of Allah did not occur in the manuscript. In the Bodleian manuscript which goes back to the 12th century or earlier, *it is Nabi Adam who gives Ravana his kingdoms and Allah taala has been substituted for the Hindu Trinity (dewata mulia raya)*. One other strong pre-Muslim element in Malay literature was a cycle of some forty tales enacted in the shadow-plays of Java, Bali, Malaya, Siam and Cambodia, whose hero is a Javanese prince Sri Panji and heroine Chandra Kirana, Moon-beam.
- G41 100 In both, however, this ambition was partially frustrated by a shared egocentricity, a neurotic self-absorption hitherto unparalleled among great writers. For Joyce as much as for Proust, *it was the 'I', the moi, with which he was ultimately concerned*: both were autobiographers for whom the objective world about them was largely subordinated to their own specialized and highly subjective mental attitudes.
- G43 91 When practical work has begun, a producer has above all to be able to give all the collaborating experts their heads when desirable, and to check them gently but firmly—that is, tactfully—when necessary. It is rather like driving a team of fine, high-mettled horses: *it is they who do the work, but, unless they are a team used to working together, they may have to be guided*. How often does an excellent conductor wish to take a passage of music at an 'effective' pace that is unsuitable in the circumstances?
- G44 10 It reached the climax of its career in the year (date), with the not inconsiderable total of performances; soon after this triumph, however, the average number of performances per year dropped sharply from about (date) to about (date), and after 1890 it disappeared almost completely from the repertoire. *It was not until (date) that the 'lost scenes' of the 'Amsterdam edition' were rediscovered and published by the grammarian, M.-J. Simonnin; not until (date) that the original Don Juan was restored to the stage at the Odéon; and even then, not until some six years later that the Corneille version was finally ousted from the Comédie Française*. The date (date), therefore, is usually taken to mark the critical turning-point in the fortunes of Molière's play.
- G44 58 The musical public of Paris in was unable to digest German opera in any form; any opera written in Germany had of necessity to be 'arranged' in the French, or,

slightly later, in the Italian tradition, if it was to succeed at all; and *it was in fact the eventual discovery that both Le Nozze di Figaro and Don Giovanni, despite their having been written by a German composer, were fundamentally Italian operas, and so might be thankfully handed over to the opera buffa, that finally established Mozart's operatic reputation in France.* The one traceable attempt to produce a Mozart opera (*Die Entführung*) in the German tradition was so disastrous and lamentable a failure that not an echo of it remains throughout the century.

G44 155 Public taste in music was guided exclusively by men of letters, and, during the whole Napoleonic era, the major dramatic critics were wont to look upon opera as their exclusive prerogative. Above all, *it was Julien-Louis Geoffroy, the feared and influential oracle of the Journal des Débats, who could make or mar a composer's reputation with a single article, although—as he thankfully admitted—music was an art which he understood no more than Morris-dancing.*

G45 178 In an earlier volume in this series, Mrs. Rice, who is Russian by birth, took as subject the Scythians. Despite chronological difficulties, *it is they who have been suggested as the link between the arts of Central Asia and the Steppes, and so ultimately with certain traits in the Scandinavian and Celtic cultures.* In his geographical history of the Vikings, Professor Arbman shows how the Rus, or the Swedes of Muscovy, traded in Black Sea ports and sent caravans into Baghdad.

G48 48 For the true naïve painter, on the other hand, there is no margin between his intention and his result: he paints to the exact limit of his vision. *It is exactly in his humble capacity to be satisfied with this that his naïvete or lack of sophistication lies.* It is exactly in this that his appeal lies. Rousseau once wrote to the mayor of his home town Laval, offerin to sell *La Bohémienne Endormie*.

G48 103 Whereupon Lautrec muttered something about middle-class stupidity, which was always prepared to 'admire an absurd gesture or a sunset.' Proust and Lautrec belonged to different worlds and *it was precisely the difference in their worlds that made Proust what he was.* He was the woman outside the window, able by the intensity of his desire and his curiosity to possess the ring.

G49 104 Their tradition is carried on by directors like Kozintsev, Romm and Heifetz. But *it is the young men who are profoundly changing all our old ideas.* The pattern began to emerge when Chukhrai made *The Forty First*, and it was consolidated

in his more recent film, *Ballad of a Soldier*.

G51 30 For the most part, this pastime has no permanent effect on the language, but occasionally, so strong is the desire to make familiar that which is strange, that a word is changed—either in whole or in part—in accordance with the fancied etymology, and the changed form is henceforth accepted. *It is a change of this kind which is often specifically intended by the use of the term 'folk etymology.'* A good example is a plant, proverbial for its bitter taste, namely wormwood.

G51 58 In fact, the word *oecern* is related to *oecer* 'a field' (modern *acre*, which has, however, become specialised in meaning), and originally referred to the produce of the fields in general. *It is not the observation of likenesses which is at fault in popular etymology, it is the fact that conclusions about the relationships of words, drawn from comparisons, happen to be erroneous.* It is not, however, necessary for a whole word to be transformed in order to satisfy the popular etymologist.

G53 83 In Britain, on the other hand, it has been assumed that the welfare of society demands the unquestioning and habitual acceptance of the supremacy of Parliament, a Parliament which cannot limit its own competence and cannot bind its successors. No doubt *it was in the seventeenth century that the decisive steps in this direction were taken*, but it would be a mistake to read into the constitutional debates of those days the modern conception of Parliamentary sovereignty which grew out of them. The truth is that all three participants in the constitutional conflicts of Stuart times in some degree accepted the notion of fundamental law and were largely ignorant of the notion of sovereignty as it was later formulated.

G53 138 It may now be safely said, with certain qualifications regarding Question Time and Adjournment Debates, that the primary business of the Commons has ceased to be the rectification of private grievances and has become the enactment of public legislation. Large and highly disciplined Parties emerged with organised followings in the country, so that *it is only on a minority of issues that the House of Commons can formulate an independent view*. Indeed, the best contemporary exponents of the constitution, like Sir Ivor Jennings, have no hesitation in holding that the real business of Parliament is to sustain government in office.

G53 145 Public interest has largely shifted away from Westminster to the Party conferences and the private conclaves of Parliamentary Parties, each of which is supported by a highly developed bureaucracy. *It is at these places, after all, that*

things really happen, that general plans of future legislation are formulated, subsequently to be embodied in election programmes. A victorious Party at an election tends to assume, often with little justification, that it has been authorised to carry out in detail the measures listed in its programme, measures conceived by Party bureaucrats, born at Party conferences and designed less to reflect the will of Members of Parliament or even that of the country at large than to appease the Party zealots.

G53 186 Now, in whatever way government may be theoretically conceived, it is in practice a matter of the adjustment of a multiplicity of private interests. If the function of an Act of Parliament is to establish general principles and rules, the details must be filled in by someone, and *it is to the civil service that the task of filling in these gaps has fallen in modern times.* Over the last half-century Parliament has perforce delegated to Ministers and to subordinate organs of the executive the task of devising the measures needed to achieve the objects of its legislation, and the measures thus devised, although they have lacked the direct consent of Parliament, have been endowed with all the force of statutes.

G54 76 The Voce di Trastevere opened a nation-wide subscription fund. *It was not until several weeks later that Tagliabue was detained by the Foggia police for simulating an offence.* He had been sweating up that snow-covered hillside, he explained, reflecting that he would not be pensioned for another fifteen years.

G60 179 Yet there are recorded, in the famous code Hammurabi, Babylonian laws relating to medical practice. *It is however from the Egyptian papyri, especially of Edwin Smith and Ebers found at Thebes and dating from about the sixteenth century B.C. that we find the first records of the practice of medicine.* These papyri show that the Egyptians shared with the most primitive medical folklore the concept of animism viz. that disease is caused by the evil influence of enemy, demon, god or even animal and that this evil spirit might be warded off by amulets, propitiated by sacrifice, and expelled by incantations.

G61 31 There is another contrast. In England it is only Oxford and Cambridge which set standards of prestige for universities. Men come and go easily between Cabinets, Embassies, Chairmanships of Boards and the Oxford and Cambridge colleges.

G61 102 The teachers in public schools and grammar schools will have a strong bias to Arts and pure science, a bias towards Oxbridge, which diminishes as one goes

down the long ladder of social status, which is not necessarily a ladder of ability or even of success. *It is to these 'customers', the advisors of students, the creators of ambition, that we have to sell a new picture of the system, as it will be, a system in which Oxbridge will have a special but not predominant place.* My last point is that to me, as a professor in a civic university, interested in the growth and government of cities, with a young family growing up in a city, the civic situation seems a peculiarly advantageous one.

G64 25 It is admitted that the search for scientific truth may be a genuine aspect of culture, and the current fashion is to praise scientists for their broadmindedness rather than call them illiterate. *Today it is the technologist who is the object of humorous deprecation.* This shows that we have not really begun to solve the problem of 'the two cultures'.

G65 38 54) The following table gives some indication of percentage allocation of investments ... *It is in the field of external resources that the greatest difficulty arises in estimating the budget of the Third Plan.* Considering foreign trade trends, the Draft Outline estimated ... the total export earnings over the Third Plan period ...

G66 89 There was no intention, declared Professor Hromadka, Dean of the Comenius Theological Faculty in Prague, in his opening address to the Conference, 'to level the organisational differences, the diversity and riches of the heritage and legacy possessed by the individual Churches and their members ... On the contrary, *it is here, among us, that our multiformity assumes a deeper meaning* ... We cannot labour for a new atmosphere in the world, in international relations, unless we form here among ourselves an internal partnership of trust and willingness to learn from one another.'

G66 167 Moreover, in the world of today it is a problem confronting Jews no less than Christians. And if the difficulties at present seem greater in the East, where the apostles of the Marxist-Leninist form of dialectical materialism openly attack what they regard as religious or superstitious survivals, the situation is hardly less serious in the West where more practical forms of materialism are in danger of undermining the very foundations of the Judeo-Christian way of life. *It is, I believe, the fact that Christians (and Jews) on both sides of 'the Curtain' face similar if not identical problems that gives special importance to this 'First All-*

Christian Peace Assembly,' and to all that went to its making and that will, it is hoped, flow from it.

G66 185 From the very earliest days infants are imbibing the implicit assumptions of the society in which they live. *It is just because of this liability on the part of young people to pick up the prejudices of their environment that the Council of Christians and Jews has always regarded the broadening of their minds and sympathies through contacts with different religious, racial and cultural groups as an essential part of its educational programme.* Here is a description of a most valuable piece of work on these lines carried out by the Leeds Branch of the Council as part of their programme and some of the reactions to which it gave rise.

G68 158 On the announcement of the primary results Sir Ferdinand Cavendish-Bentinck could justifiably claim an outstanding triumph, but this was only the first stage of the election. The principle of Kenya's new Constitution established at Lancaster House was the common roll, so *it would be the mass African vote that would prove decisive.* Would Sir Ferdinand's be a Pyrrhic victory?

G72 84 The National Certificate Courses at Ordinary and Higher levels have provided the training for many technicians and an avenue to full professional status for many students. *It is in these courses that the high rate of failures has attracted most attention.* The White Paper proposes that the Ordinary National Certificate course, at present a three year course, should become a two year course i. e. the length required now of students who are exempt from the first year because they have the appropriate passes at Ordinary level in the General Certificate of Education.

G74 181 Briefly, the answer is two-fold: political vigilance and military professionalism. *It is the imperfectly professional army and the careless statesmen or power vacuum which constitute the ingredients of military intervention.* But the successful containment of military power within its proper sphere has never been achieved without difficulty.

76 192 When General de Gaulle came back to power two and a half years ago, there was a general wave of optimism about his chances of bringing the tragic problem of Algeria to a settlement. Algeria was in the forefront of every Frenchman's mind at that time, because *it was a crisis in Algiers that brought about the appeal to de Gaulle to return.*

77 81 These digressions have seemed necessary because, despite the title of the Obscene Publications Bill, the connections between obscenity and pornography are both tenuous and intermittent. In Latin literature such writers as Juvenal and Martial used the complete obscene vocabulary without apparently being considered pornographic; we do not know what vocabulary Elephantis and her colleagues employed, but for her contemporaries *it was the subject matter, not the language, which made her books reprehensible*. Conversely, to the best of my recollection, The Memoirs of Fanny Hill (one of the few masterpieces of English pornography) does not use a single obscene term.

Wh-Clefts (72 examples)

A04 142 He was afraid lest the decisions would mean the labelling of the party at a future General Election that it did not 'care about the defence and security of our country.' After saying that Mr. Khrushchov believed Russian power to retaliate had stopped a U.S. attack upon the Soviet Union, Mr. Gaitskell said that *what stopped the Russians in the last resort from aggressive nuclear war was the certainty that they would be annihilated*. Mr. Gaitskell added that agreement with the unilateralists was not possible, though he respected their views.

A08 165 They are also the answer to those critics of Conroy who complained that he slowed up the attack. *What they meant, of course, was that Conroy did not belong to the school that bash on regardless and hope for the best*. Smith, who played outside Conroy on both wings for England, always says that nobody else could place the ball so perfectly to create an opening.

A12 98 Anyone who went into unskilled work, he said, went in at the rate for the job, regardless of age. *What many parents did not seem to realise was that a relatively high wage now might be a rather poor one in after years*. Britain, said the Duke, could not hope to compete in foreign markets if industry went on using unskilled labour.

A19 154 I'll do no work for weeks and then write solidly for 12 hours. I think *what I'm really seeking all the time is the source of Original Sin in myself*. Logue leaped to his feet at this heresy and shouted: Original Sin! What are you talking about?

A19 214 'But even Auden can't make a living just writing poetry,' said Roethke. 'I doubt if anybody does, except maybe Robert Frost'. Let's face it, poems will never be as popular as football coupons, and *what America offers is just bigger subsidies*. As characters, poets range from rhyming layabouts to saintly travellers who have embarked on the greatest journey of all: the journey into the mind and spirit of man.

A25 128 Here, a company is formed to exploit the actor's services over the next five years. In essence *what happens is that the drop in value of the shares in this company (at the end of the five years they are worthless) is offset against the actor's earnings over the period*. How many of our leading actors are anxiously waiting for April 17 to find out whether they will be able to continue their forward strip?

A26 66 He shrugged: 'Might do one good thing—prove once and for all there's no God up there.' *What is taken for granted, both in America and in Russia, is that there would be no sputniks if there were no military rockets*. Which brings me to the gloomiest aspect of my journey: the spectacle of two great peoples getting ready to massacre each other.

A26 82 I switched to electronics, and it was the same. I sat down to figure out what they couldn't use for war, and *what I came up with was birds*. Add to this a still potent distrust of foreign entanglements.

B01 64 Britain and the U.S. , which have problems with their balances, will gain some immediate help. *What it means in practical terms is that our exports to Germany will now be a little cheaper for Germans to buy, while the goods which Germany exports will be made a little dearer*. Booming Germany is deliberately encouraging more imports as a means to curb rising prices at home.

B01 87 If, however, in addition to her new good-neighbour gesture, Germany takes a really big share in giving aid to underdeveloped nations, the world outlook will be brighter. *What gives rise to optimism is the sign that Germany and the other leading Western nations are at long last moving towards a solution of currency problems by co-operation*.

B02 26 But that might give President de Gaulle a convenient excuse for keeping out of talks if he still thought this was not the time to start them. *What seems certain*

is that those who advocate putting off any approach until Mr Krushchev gives evidence of a change of heart (whatever that may mean) would have us run risks greater than the West ought to run and greater than President Kennedy's most influential advisers seem disposed to face. The real question is what we should put to the Soviet Government as a basis for talks :and that means working out what we know to be the essential interests of the West in Berlin and what we suppose that the Soviet Government may now be after.

B03 42 But discussion on current points of east-west conflict could not go much farther than, for example, the truism that policies of menace and mutual disarmament cannot be followed together. *What the council has done and it is an achievement is to make religious contact across the greatest political barrier in what is not yet a unitary world.* In the words of one Russian delegate, older churches like his own have personally discovered younger churches for the first time.

B05 67 They have developed a new system whereby completely untrained workers can be taught their trade by means of tape recordings and television. *What happens is that the unskilled worker is processed, by high-speed listening to recorded instructions on how to do the job coupled with explanatory TV pictures, into becoming a highly skilled, obedient craftsman in no time at all.* Not only can the raw human mind be technically equipped very quickly to do one set of skilled manufacturing processes in one trade but, by being given another of the new audio-TV training techniques, he can be switched to a different industry if he just gives in and listens and looks.

B05 173 I may well have felt that I was hell-bound under a strict Presbyterian upbringing and a possible reprieve might come through the sugary sentiments of Abou Ben Adhem. *What happened in the jingly-jungly jingle was this: Abou Ben Adhem (may his tribe increase!) Awoke one night from a deep dream of peace And saw, within the moonlight in his room, Making it rich, and like a lily in bloom, An Angel writing in a book of gold. To cut a long story short the Angel got on well with Abou and wrote his name at the top of the book of gold.*

B09 147 In spite of post-Moss Side, post-Oswestry, and cosy Liberal optimism it will be some years yet before the Liberal Party can form a Government, while it is obvious that the unique event of a Labour majority in the Commons is unlikely to be repeated. But *what we need is not a Lib. -Lab. pact but a new party; not*

coalition but coalescence. Is it too much to hope that the Radicals, now sprinkled in all three parties, may one day be united and that the Liberal Party may find itself the anchor of a new radical alliance?

B09 173 I am unable to find this reference in the report which is hardly surprising as it was published in 1959 and its latest reference to costs is in the financial year 1957-58. *What the Hinchliffe Report does say on page 27, paragraph 63, is : These figures do not support the general belief that the cost of the pharmaceutical service is increasing at a much faster rate than that of other branches of the National Health Service or that it is absorbing an increasing share in the total cost of the Service.*

B11 203 This was never my intention and such an interpretation is possible because my remarks were of necessity condensed. *What I intended to imply was that doctors often prefer not to be used solely in a consultative capacity in recommending treatment for patients to their general practitioners because they prefer to prescribe such treatment themselves, to know that it has been supplied and then to follow up their patients by seeing them again.* This situation can be realized by the use of E.C. 10(H.P) forms written by hospital doctors and dispensed by retail pharmacists.

B12 56 Nor is his undoubted success entirely accountable in terms of his personal charm, great though that is, nor of the presence of the Russian Trade Fair. *What in fact Major Gagarin seems to have done is to have shown us how much we want to like the Russians, in a spirit of genuine neighbourliness.* This, and the fact that British visitors to Russia usually find a reciprocal warmth of welcome there, is surely a portent worth noting by the political leaders on both sides.

B15 106 There can be no real objection to these proposals, in principle. *What should concern us is how this policy is to be administered, and whether in fact it can be administered without racial discrimination.* Since the majority of immigrants to-day are coloured, it will be difficult to avoid the suspicion of discrimination.

B16 32 The darlings of democracy today are the men who, long after the Communists have shown their true colours, have handed country after country to them: Benes surrendering Czechoslovakia, Roosevelt giving them half of Europe, Truman and Attlee abandoning China to its mild agrarian reformers. Might I suggest that Moscow knows that Kerensky has been one of its most unhesitating and de-

terminated enemies for 44 years, and *what it is really looking for in Persia is not Kerensky (nor Mikhailovich, nor Chiang) but a nice Western-style statesman with half-a-round-table-full of crypto-Communist advisers?*

B16 47 It reads : In so far as the United States has hitherto been the protector of the regime, the people tend to be emotionally anti-Western ... *What I had written was : ... the opposition tend to be emotionally anti-Western.* What proportion of the people belongs to the opposition is a matter of opinion.

B20 190 The man who captures it will go down in history as one of the greatest of mortals. *What the great masses of ordinary people in the world desire most of all is the certain prospect of peace for as long ahead as possible.* No one can blame Harold Macmillan for trying to reach the elusive goal.

B21 114 They can laugh at their own weaknesses, like the belief that it is impossible to eat in a restaurant without gipsy music, although the gipsies have disappeared long since. *What I found depressing was the insistence that all the many good things in the country were due only to socialism and the Party and would not otherwise exist, together with fantastic ignorance of the western world or refusal to believe what did not suit the theory.* To give a couple of instances that stuck in my mind ...

B21 184 This was, of course, untrue, the negotiation was a flat demand for a Soviet officer to be brought to whom they would talk and with whom they would, no doubt, establish their identity. *What is at stake, in fact, is whether the West recognises Herr Ulbricht or Mr Krushev as responsible for East Berlin.* The West has no intention of recognising Herr Ulbricht, as it has made very clear to the Russians in public and private conversations.

B25 125 It took place on a part of the A that has a dual carriageway which the people of Sidcup are still hoping will be extended into this urban district so the need for a road improvement of that nature cannot be argued in this case. *What is alarming is not only that this sort of accident can still happen with dual carriageways, but that there could so easily have been other fatal accidents within the urban district over the week-end.* A number of brushes between traffic was reported to the police, several of them causing minor injury.

- C07 10 Lautrec's liking for whores and dancers and singers and acrobats as subjects was, of course, a perfectly commonplace taste among artists of his time. *What is singular about his use of them is that no other artist, of his time or any other, has painted them so directly, intimately and pertinently.* He doesn't, on the one hand, use them as symbols, pegs for a moral or aesthetic attitude, as the young Picasso does (to take one example among many); and on the other hand, he doesn't use them only for the way they look, like Degas, whose dancers are more or less interchangeable with his laundrywomen—the same breed with a different set of gestures.
- C08 142 The truth is that at least a dozen selections of equal size, equally good and equally representative, could have been made from the same source. *What matters is that the editor has given us the essence of Wellington's genius—his clarity, his good sense, his powers of observation, his understanding of human nature, his dry irony, his wonderful balance and foresight.* It is like offering the reader a small parcel of a superb cellar; it is all there for his buying if he wants more.
- C10 163 He never got up much speed, and made few turns. *What he did do was to fix us with a basilisk stare, make odd pointing gestures and keep improvising for about twenty minutes.* A polite attempt to drive him offstage with a burst of applause only spurred him to go on and on.
- C12 55 Nothing could be more remote from Cavafy than any of these. *What is in fact the case is that he was concerned with a view of a Greek's place in history, a view which was peculiarly his own and which has been found by his contemporaries and successors in the Greek tradition peculiarly true and enlightening.* It is a view taken from 'a slight angle to the universe', but is none the less accurate for that.
- C13 172 Phoenix is itself a miscellany, unplanned, yet unified as no other miscellany could be, by the personality of Lawrence himself. Some of the things it contains are of rare quality, some interesting for what they add to our understanding of Lawrence's 'philosophy', some are comparatively trivial pieces; but *what matters even more than their individual merits is the cumulative effect which they achieve when brought together in this way.* The sum even of the novels and poems is greater than the parts, but the existence of a collective meaning, subtly influenced by the presence of the author (which is always felt in Lawrence's work), can be more easily perceived in the sum of Phoenix.

- E03 114 Almost certainly they are of carbon steel—good enough for a hand brace, maybe, but you would be lucky to use them more than once in a power tool. No, *what you need are high-speed-steel drills, more expensive in first cost but cheaper in the long run*; these can tackle most jobbing work, but you will want yet another type of drill if there is glass, concrete or masonry to be drilled. A good high-speed drill will have the letters HS and the diameter stamped on the shank.
- E08 156 It becomes quite a problem to know how to keep everything in good condition and it is more important than ever to use the ventilators as much as possible whenever outdoor conditions are favourable. *What is essential is to maintain a free circulation of air around the plants and to keep the atmosphere reasonably dry*. Schizanthus and calceolarias need very careful watering at this time of the year.
- E18 126 Cooling water in vast quantities will be discharged into Southampton Water but except for being warmed it will be unchanged. *What we can and do object to, however carefully 'landscaped' and however beautifully designed this power station may be, is the fact that we shall be able to see it from all parts of the Solent*. This may seem a slight objection but it is a valid one.
- E38 36 Are their present services so inadequate that it is necessary for the N.F.U. to step in and improve matters? *What triggered off the move was a report, published last year, on the marketing of herbage seed*. The N.F.U. was rightly concerned with ensuring that British seeds had at least an equal share in their home market with foreign seeds, and there is much, though not everything, to be said for a standardised product.
- F01 148 For the moment we need not concern ourselves with these last questions. *What we have to grasp is that there are patterns of brain-activity of different kinds*. There are patterns which result directly from processes such as seeing, hearing, etc., and that which is experienced as the shape of the cloud (but not yet identified as such).
- F13 76 The responsibility of the advocate in court rests upon the importance of every witness's honest recollection being fully tested. When a man comes before a court charged with, say, driving dangerously, *what it really means is that 'in the opinion of a number of witnesses, whom you will see and hear, he was driving dangerously.'*

- F15 14 It is not a question to be decided on nationalistic or political party feelings, and no reasonably objective opinion, one way or the other, is likely to be arrived at without going through the process known to the writing world as beginning with Adam and Eve. *What that means in this case is going back to the war years, when the Governments of Belgium, Luxembourg and Holland were in exile in London and had every reason to be so concerned about the precariousness of their post-war prospects that they organised a careful experts' study of the subject.* The outcome, based on the realisation that their chance of economic recovery and their ability to make themselves heard in international politics was desperately poor if taken singly, and decidedly better if they could act in unison, was the agreement to join their three countries in a union to be called Benelux—in which we can now recognise the beginning of all endeavours to unify Western Europe.
- F21 36 On the contrary it is a time for constantly taking up new pastimes, new interests. *What must be dropped is those physical leisure-time exercises taken up in youth and now inevitably being performed with diminishing success.* A man, it has been well said, whose enjoyment consists of constant reminders that he is not as young as he was should take medical advice immediately.
- F27 24 Men of the period, both humble and gentle, accepted a stratified society: *what they resented was the abuse of official or social position,* and this is precisely the attitude which the ballads echo, with their detailed catalogue of the crimes of men like the sheriff of Nottingham and the Abbot of St. Mary's. One should not expect popular literature to concentrate its attack on the manorial system or the inconvenience of villein status, because the peasants themselves did not see their grievances in economic or systematic terms: they saw them rather in terms of the personal viciousness of individual lords.
- F31 41 There was no fracture and all that had been needed was the simplest treatment. *What they had done was to take a three-inch elastic adhesive bandage and apply it carefully but firmly from below upwards, so that it supported the torn outer ligament of her ankle.* I encouraged her to try walking on it now that it was safely strapped up.
- F34 99 At that age people tend to become somewhat set in their ways. *What is remarkable is not so much that the Royal Academy should have remained distinctly academic, but that it should have shown so much tolerance as it has to the younger*

men.

F35 57 The event is celebrated each Sunday in summer when the story is re-enacted by a piper and boys, the latter disguised as mice. Unfortunately, modern research tends to discredit the legend, claiming that *what really happened was a visit from a labour agent who attracted many local young men away to Bohemia, with the promise of good wages.*

40 111 It is necessary, to avoid confusing the issue, to ignore some of the extreme examples of deleterious sounds, those that make telephone operators faint or the jingling of a bunch of keys that sends a mouse into something approaching hysterics. *What is at least as interesting is the way inventors seem to have chosen, probably intuitively, a combination of explosive and aggressive sounds as warning signals to be used on automobiles.* Apart from the purely explosive sounds, those that stir most animals to rapid action are the snarls, growls, barks or long drawn-out roars of predators or rivals.

G14 28 These call themselves democrats, but as they have never yet got full control of the footplate, nobody knows what their large claims amount to. *What arouses the indignation of the honest satirist is not, unless the man is a prig, the fact that people in positions of power or influence behave idiotically, or even that they behave wickedly.* It is that they conspire successfully to impose upon the public a picture of themselves as so very, very deep-thinking, sagacious, honest and well-intentioned.

G15 51 The well-worn antithesis between Judaism as the religion of Justice and Christianity as the religion of Love does not, in Ahad Ha-Am's opinion, go to the root of the matter. *'What essentially distinguishes Judaism from other religions is its absolute determination to make the religious and moral consciousness independent of any definite human form, and to attach it without any mediating term to an abstract, incorporeal ideal.'* Hence the Christian idea of a divine-human being, who mediates between God and man, is one which Judaism can never accept ...

G36 13 The aliens, then, are on the airfield all right, but their space-ship is sinking into a muddy heelprint or whatever. Apart from the effects of awe and amazement produced by the description of the pulpy monsters and so on, *what we have here is a strong puzzle interest that is widespread in science fiction as a minor aspect and*

not uncommonly central, as in this case. I have already mentioned the biological puzzle—problems of determining an alien life-cycle and the like—as an important sub-category; another involves the question of finding the weak point in some apparently invulnerable monster or hostile alien or badly behaved human artifact of the robot sort.

G36 99 Going easy on the puritanism would be a commendable resolve, and so would a decision to drop sex altogether where it is not essential rather than to decorate a planetary survey or alien invasion with a perfunctory love interest presented in terms borrowed from the tough school or the novelette. *What will certainly not do is any notion of turning out a science-fiction love story.* In the as yet unlikely event of this being well done, the science fiction part would be blotted out, reduced to irritating background noise—a dozen Venusian swamp-lillies being delivered to the heroine's apartment, and so forth.

G37 41 From his knowledge of war as it really is the poet may start again towards a wider vision of it and try to see it in a fuller perspective without reverting to the old abstractions and falsities. It is impossible to present its illimitable chaos, but *what counts is the poet's selection from it of what really strikes or stirs him ...*

G38 122 We all know that certain passages in good fiction and good poetry are used by some readers, chiefly schoolboys, as pornography; and now that Lawrence is coming out in paperbacks, the pictures on their covers and the company they keep on the station bookstalls show very clearly what sort of sales, and therefore what sort of reading, the booksellers anticipate. We must, therefore, say that *what damns a book is not the existence of bad readings but the absence of good ones.* Ideally, we should like to define a good book as one which 'permits, invites, or compels' good reading.

G44 52 If Moliere's heroic seducer was unfortunate in the manner of his reception by the Parisian audience, his operatic counterpart was scarcely less so; and the trials and tribulations of Don Giovanni at the Grand Opera furnish an admirable illustration of the obdurate tenacity of French musical conventions, which, in the post-revolutionary period, were certainly as rigid as those of the Comedie Francaise, and even more fettering to would-be dramatists of the new generation. In this brief study, however, *what interests us is not the direct significance of these musical conventions in themselves, but their indirect influence upon the fate of*

Moliere's Don Juan. The musical public of Paris in was unable to digest German opera in any form; any opera written in Germany had of necessity to be 'arranged' in the French, or, slightly later, in the Italian tradition, if it was to succeed at all.

G44 163 The story of the resplendent premiere, the gradual disintegration and eventual catastrophic débâcle of this first French production of Don Giovanni can be followed in detail through the reviews in the contemporary press. *What appears evident from the various comments which have survived is that Kalkbrenner's manipulations of the score had put all the critics except Geoffroy in a quandary.* Geoffroy's position was simple and unassailable.

G48 165 Then, in the small commercial galleries, a desultory collection, out of the tourist season, of Roman and other Italian artists fighting their battle against what is expected of them or giving themselves up to an illusory affair with some faded beauty-spot, and coming out of it rather worse than such ill-advised lovers elsewhere. *What is instructive is to see the three aspects of modern art—realist, abstract, and that curious cabalistic art of symbolism and fantasy mixed that has no tidy name—in a new setting and a new light.* Certain things become very clear.

G53 43 In being pragmatic about his priorities he will rightly emphasise on the one hand, for example, the political advance of Somalia, while arguing on the other that a more complex set of constitutional checks and balances is required in Northern Rhodesia. The jibe of Mr. Mboya in attacking the Lancaster House Conference that *what Somalia required were settlers* was double-edged. Settlers might have retarded Somalia's political progress, but they would have given it a much better standard of living.

G53 43 Here we should look to a judge or a retired and senior Treasury official, or to Parliament, to provide such services if they are required. *What matters is that Denmark appears to have found a way of satisfying what is prima facie a legitimate public demand for protection against administrative abuse without either paralysing administration or diminishing the dignity and independence of the judiciary.* This, to say the least, is a constitutional example worthy of scrutiny in the context of other political and social circumstances which, however, include the tendency towards ever-increasing administration noted by the advocates of the Ombudsman in post-war Denmark.

G58 119 For we, in the universities, hope that they will see the problems of here-and-now- whether they are the problems of personal conduct, of public affairs, of art and literature, of science and its applications—illuminated by the studies of their university years. In other words, *what the student needs from the university is not just a little (or even just a great deal) more competence in the subjects he has studied at school; not just to have a few rough edges knocked off his mind; not just to learn more elaborate intellectual skills; not what, in the modern idiom, is called 'know-how'.* He is going to be a member for three or four years of a society which has its own characteristic way of life.

G60 52 Edward Shils and Daniel Bell both write about the 'End of Ideology', but not very convincingly. *What they really mean is the end of the appeal of communist ideas to the intellectuals.* I believe that we can still see pervasive influences of certain kinds of ideology in American thinking.

G62 109 Moreover, their use of manpower per beast or per acre is very high. *What is impressive is the enthusiasm and thoroughness with which they carry out their systems:* grooming of cows, attention to their feet, feeding of calves, detailed keeping of farm records ...

G62 112 ...But I should like to end by saying that *what impressed us most was the warmth of our welcome.* As far as we could learn, we were, in every case, the first English people to visit the farm.

G63 31 So far I have been much concerned to rebut the strong suggestion that what might be described as the carrying out of an intention could be known without observation, but now I want to return to a weaker suggestion which was shelved at an earlier stage. This is the suggestion that *what we know without observation are our intentions.* One might perhaps concede that neither the driver in my example, nor the man writing on the blackboard in Miss Anscombe's, could know without observation that their respective intentions were actually being carried out, but one might also claim that in both cases the persons concerned would know what they intended to do and would know this without observation and quite independently of what actually happened.

G64 99 Our error is not in training scientists who are unaware of the classical outlook: it is in training them in all sorts of assumptions which are still unconsciously derived from it. *What we need, to produce scientists who are also human, is*

something far more fundamental than a Departmental Committee on Syllabus Revision on which schoolmasters and industrialists as well as university dons are represented (although that would be a practical first step which is already long overdue). We need a radical revolution in our whole outlook.

G64 161 Those who have absorbed the atmosphere of scientific culture find those outside it alarming because they appear to be willing to attach more validity to their fundamental myths than to evidence. *What the new men want—and will have, sooner or later—is a public system which bases authority always on declared evidence that the good of persons is demonstrably being served.*

G65 163 Poverty, unemployment and illiteracy have yet to be mastered completely; and the common man cannot, in general, feel relaxed under the umbrella of the welfare state. Nevertheless *what is surprising is not that planning has achieved so little in its first ten years, but that it has achieved so much in so short a time in a country which inherited problems created by centuries of foreign rule.* Before the war, India was almost completely dependent on foreign countries for the most elementary articles of consumption—from needles to locomotives, and from tooth paste to heavy chemicals.

G67 104 The contending approaches I've just cited (I could add some more, extending to scholastic disputes about which techniques ought to be applied in—entirely hypothetical—investigations) of course contend mainly in the minds of the disputants. *What makes so many of these debates so sterile is that the participants either cannot or will not see that they occupy vantage points of a very restricted sort; they seem to think that, like so many intellectual colossuses (sic), they straddle the globe.* The more one looks at this, the more one feels that the thing which British sociologists need is to consider the implications of Weber's work for their own.

G73 7 Nearly twenty years ago, when D.S. Senanayake asked me to prepare a Draft Constitution for consideration by the Ceylonese Ministers, I asked him what sort of Constitution he wanted. He replied that he was not very concerned with the details, because *what he wanted was a transfer of power from British to Ceylonese Ministers.* I have heard that sort of remark several times since.

G73 12 First, nobody can transfer power, except in a purely legal sense. *What is transferred is legal authority, and legal authority does not necessarily confer power.*

If you have legal authority to knock a man down, you still have to knock him down; and he may prefer to knock you down.

G73 42 Nationalists are nearly always impatient, and they often think that the British Government is being deliberately slow and evasive. But *what the Colonial Office really tries to do is to glide so gently from colonial rule to independence that the machinery of government will go on ticking over as if no fundamental change had taken place.* Some of the Nigerian leaders came to London ...

G73 109 That is a common experience; but a good deal of knowledge has been accumulated over the past twenty years. *What I am sure about is that all the problems which can reasonably be foreseen ought to be solved—in so far as they ever can be solved—before the transfer of authority takes place.* In other words, a detailed and permanent Constitution ought to be carefully worked out beforehand.

Reverse *Wh*-Clefts (38 examples)

A04 128 It was far better for a weapon used for retaliatory purposes to be under the sea rather than on land. *This was why the Labour Party did not think it right to oppose the Polaris depot ship.* The party agreed that it was unwise to locate the base in the Holy Loch, only 30 miles from Glasgow, a city with two million people.

A16 69 Ordinary shares of Sopers of Harrow, Middx, the store firm, are worth 10,0s. each. *That is what Debenhams, which already owns most of those in issue, are offering for the remaining 4,135.*

A22 236 Cheque in Scotland! Excuse the play on words, but *that's what it could amount to where Spurs and Chelsea are concerned.* Bill Nicholson, Spurs boss, has money to spend to maintain Tottenham's Double Top League and Cup glamour.

A37 79 A photograph of this occupies pride of place on his desk in the Fleetwood office. 'Since she was very young,' says her proud father, 'she has always taken a great interest in the business, and *that is why I chose her for the company's trade name.* I have confidence in her ability.'

A39 84 Sometimes when you listen to a record, the backing, however unobtrusive it may be, attracts your attention more than the singer. I mean no slight on Mr. Gary Miller when I confess that *this is what happened when I heard his record 'Dream Harbour' (Pye, N.15338).* The accompaniment, a soft, oriental rhythm,

came through entrancingly.

B01 181 Only a man wrapped in the impenetrable cocoon of what he regards as a divine mission could have spoken of apartheid as a policy of good neighbourliness. We may be sure that he is not being hypocritical. *That is what he really believes.* A good neighbour to those Africans who, under apartheid, will be forced back to their tribal reserves with no prospect but a cramped and primitive existence.

B14 77 And if he needs spectacles, when in hospital, he gets them free. And yet a widow, whose pension, for which her husband paid, is wiped out because she works for a living wage, will now have to pay 12 s. 6 d. for each lens in her spectacles, and 17 s. 8 d. for the frames. *This is what the Minister proposes.* The truth is that you can't make sense out of small private charges under a vastly expensive public scheme.

B14 84 But, at least, it ought to be a Minister's duty to refrain from doing positive harm just to collect a token tribute to the total tally. And *social harm, I fear, is what two of the proposed changes are going to achieve.* For example, from March 1, each item on a National Health Service prescription is to cost 2 s.

B17 104 This was, of course, a reference to the directive by the TUC General Council before the actual expulsion of the union, asking that Mr. Foulkes should resign his office and submit himself again to the members for re-election. There can be no doubt, however, that *readmission is what the members, or at least the vast majority of members, of the ETU want.* But apart from whether or not this actually comes about, a heavy blow has been struck against Communist influence, one that should, and could, have been struck long ago.

C07 36 He must have suffered not only from knowing what a monster he was to look at, but also from the uselessness to himself of his distorted body. *This perhaps is what gave him a fascination with bodies that were agile, bodies that could do what was asked of them, and bodies that others wanted to use.* At the same time, he needed to reassure himself about his own deformity with his consciousness that these bodies also would in time become, as his had, useless and hideous and unwanted, and that to would become so through the very exploitation of their desirability.

C10 160 There can be no doubt that, like Romeo Coates, he believes utterly in his mission. 'Dance inspires him ceaselessly to strive higher and higher towards the shining pinnacle of perfection that is the goal of every Artiste.' *Kathak, with its swift spins, is what bedizened boys used to dance before Mogul Emperors.* Mr. Kumar rashly did it stripped to the waist, his long hair arranged in an untidy bird's-nest.

C17 161 It is strange that a Hollywood actor should get the idea for a film in a New York students' loft on January 14, 1957, and a few months later, with money borrowed and money donated after a TV interview, make this film in the streets of that city and then fail to find anyone in the United States who would show it. *That is why John Cassavetes came to England to find someone who would take a risk on something new.* It was the directors of newly-constructed British Lion, who have got faith in fresh faces, talent, ideas and letting people try them out, who saw 'Shadows' one evening and immediately offered Cassavetes the money for world distribution rights.

D16 55 It's for families to decide what kind of houses they want and what kind of education is best for their children. *This is what the modern State usually forgets.* The Catholic Church always remembers. Hence all the quarrels between the Church and State.

D16 172 Christ, of course, did no such thing. What had He told them? 'Render to Caesar the things that are Caesar's and to God the things that are God's.' *That's what He had said.*

D16 175 *That's what he expects.*

D16 176 *...if that's what happened to Christ Himself the priest is not surprised it should also happen to him.*

E05 110 Stainless steel from Sweden and the English Lake District: *this is what we found at Pearce's, jewellers of Market Pl, Leicester.* The coffee-jug, designed by Gense of Sweden, costs 7s 6d and the three-legged dish, hand-beaten by the Keswick School of Art, is 19s 6d. Lanterns outside the front door, we are told, are fast increasing in popularity. This we confirmed whilst visiting Jack English at his shop in London Rd, Leicester.

EO5 127 Bohemian crystal chandeliers, magnificent and resplendent, to grace alike the baronial hall or 'semi-det', are hung in a superb exhibition at Morgan Squires in Market St, Leicester. Crown Theresian chandeliers are known throughout the world- Maria Theresia, Empress of Austria and the mother of Marie Antoinette, decorated her palace with their like and *this is how the name originated*. They are made in Czechoslovakia and distributed in this country by Homeshades of Baker St, London W1.

E22 47 It was not easy for anyone to believe at the time—about a year ago—when the noted scholar A.J.P. Taylor returned from a visit to Hungary, and wrote an encouraging report, that he was not misled while there. *This is what he wrote*: 'the solid unmistakable fact is that Hungarians are now pretty well off: I have never seen a greater display of foodstuffs ...there are clothes in every quality, from multiple stores to elegant private shops- in the provinces too: the days of hardship are over'.

E35 68 It should be apparent from these remarks that very careful thought and planning should be undertaken before carrying out any type of advertising. *Quality is what counts*. As has already been stated, it is not always the size of space which dominates but how the space which is bought, can be used to the best advantage.

F01 173 Before going further we should try to face what is an almost inevitable difficulty for anyone approaching the study of mind for the first time. It is the tendency to get things the wrong way round. As a psychologist I am constantly encountering this tendency in friends and acquaintances. They think there is something inevitably 'queer' about psychology and this feeling of queerness usually boils down to a quite mistaken belief that the psychologist first looks into his own mind and then interprets other people's minds by what he has found in his own. *This is what I mean by 'getting things the wrong way round'*. He is far more likely to find out about how his own mind works by looking at other people's.

F13 85 If a witness can be persuaded by an advocate in cross-examination that his honest, preconceived opinion must have been wrong, then that witness's side of the case suffers a major blow. *That is why the defending advocate is watching you at this moment with such intensity*. He is trying to read your mind, to understand your prejudices, to assess your qualities of reason and of reasonableness.

F17 10 The mother's face was drawn with anxiety. 'It's my little girl, doctor,' she said indicating the fair-haired child sitting by her side. 'I'm desperately worried about her. I think she's got cancer.' The doctor showed no emotion. 'And what makes you think that?' he asked. 'Well,' said the mother, 'she's developed a lump in her chest. It's getting bigger, too. *That is how cancer starts, isn't it?*' 'How old is the child?' asked the doctor.

F27 94 And from what class were the sheriffs and justices othe ballads chosen, if not from among the knights? The fact is that the knights as a class are not treated consistently in the ballads, *which in my submission is what we should expect*. The commons had no animus against social rank as such: what they resented was the lordship of unjust men and their corrupt practices.

F27 151 The outlaws were not always poor men, but the poor man did not demand that. He demanded kindness, good lordship to engage his fidelity, and *this is what the outlaw gave*. It is the theme of Robin Hood's famous advice: But loke ye do no husbonde harm, That tilleth with his ploughe.

G07 71 She had dreaded still more that he might return to England, overshadowing her again with spiritual and social peril. But *this kind of return was what she could never have foreseen* . . . that he should come back not voluble but silent, not beautiful but defaced, not in obloquy but with his praises ringing!

G14 58 It is not the mere fools that the confidence men successfully delude. It is, in their pregnant phrase, the 'larceny in the blood' of the victim which results in his victimization. And *that was how Hitler operated*—exploiting and using as his leverage the 'larceny in the blood' of innumerable politicians in every country who wanted to believe that here was a man who really had found a way of making diamonds out of plastics; a way, that is to say, of making a quick profit out of an illicit sale of the Western soul.

G15 153 He realised, however, as not all Zionists did in those days, that there was an important difference between 'the establishment in Palestine of a national home for the Jewish people', *which was what the British Government undertook to support*, and 'the re-establishment of Palestine as the national home of the Jewish people', which was the formula suggested on the Zionist side.

- G21 155 Make your way along the Backs on a May morning to the Wilderness, penetrate passages and archways, cross bridges and gaze again and again at the Great Court of Trinity: *this, believe me, is what education means, real education, for through appreciating the beauty of things you come in time to appreciate the beauty of ideas.*
- G34 89 Yet if, because of his broad generalizations and his imperviousness to tinsel compliments, we used to think him unworldly, we were at once overestimating and underestimating him. For he has shown—and *it is why the Rede Lecture has such an authoritative ring*—a fine grasp of the realities of power. It is one reason, too, why in his novels the pictures of closed societies, clubs or departments are so horribly accurate.
- G37 43 ... *This is what Georg Trakl, who died on the eastern front in December , does in 'Im Osten' ('On the Eastern Front').* He applies to the whole shapeless panorama of battle his gift for images which form a centre for a host of associations and must be taken at their full value as each appears:
- G37 59 Yet it was not impossible for a fighting man to let his vision pierce beyond the actual carnage and to divine with an apocalyptic clairvoyance its meaning in the scheme of things. *This was what Isaac Rosenberg did.* In the British army he had little in common with his fellow poets. They were officers; he was a private soldier.
- G37 112 The soldier has to adjust his mind to death. He does so by treating it as nothing unusual, and in his topsy-turvy world he is not wrong. *This note of superficial detachment is what Guillaume Apollinaire catches in 'Exercice'.*
- G52 133 Now most experts are willing collaborators, but the danger with all experts is that they are often not content to give of their best but insist on valuing their own contribution higher than that of other experts: think of the brilliant designer Inigo Jones and Ben Jonson's not unreasonable attack on his conceit. *That is where the Mahler or Diaghilev is invaluable.* Cochran, who checked every bit of material used in his shows (like Bernard Delfont now), was always there to appeal to, and was always watching from the background ready to step tactfully in to prevent trouble.

G57 111 It would be disastrous if the law had no firmer basis than the emotions of the majority—if dispassionate reason, knowledge, and common-sense were not also allowed a voice in its determination. Yet *this is precisely what has happened in the realm of private morality which is concerned with sexual relation*. Homosexual practices between men in private are deemed to be criminal (but not lesbianism, fornication, or adultery) simply because such practices arouse in the 'reasonable man' feelings of reprobation so strong as to demand expression in repressive statutes.

G57 135 He recognizes that a morality based upon the consensus of a majority, even if that consensus is one mainly of feeling, is essentially a democratic notion, and that democracy means the running of risks which are inseparable from majority rule. But he insists that loyalty to democratic principles does not require us to maximize these risks, '*yet this is what we shall do if we mount the man in the street on top of the Clapham omnibus and tell him that if only he feels sick enough about what other people do in private to demand its suppression by law, no theoretical criticism can be made of his demand*'. And in this connection it is well to remember the adventitious and irrelevant means by which such sickness can be induced—the propaganda and pressures of many dubious kinds which can build up artificial emotions of reprobation to the point where they have to find expression, and may do so through the law.

G59 57 To his instinctive behaviour on that occasion we can, in part, attribute the development of the Labour Party within the Parliamentary system instead of outside it, at a time when Left-wing movements throughout Europe became emigre groups within their own countries. Holding the ring—for *this is what such conduct is*—is not confined to strict constitutional questions. In Sir Harold Nicolson's biography, there are many examples of George's anxiety that the dominant party or even interest should not, so far as it was within his power to influence decisions, ride roughshod over the rights of any of his people.

G73 27 A host of public servants, civil and military, have to obey orders; even then, government will not be efficient unless the people as a whole accept leadership loyally and enthusiastically. *That is why the transfer of legal authority from British to Asian or African hands has been done as slowly and as cautiously as political conditions make possible*. Long before the example of the Congo, we learned in

India in that it is possible to move too quickly; and in India there was no question of the public services breaking down because of the failure to obey orders.

2 The Spoken Data

In this part of the appendix appear the examples taken from the Survey of English Usage (SEU) data, appearing as Svartvik and Quirk [1980], and examples collected by the author during the course of the study (1987–1989). A superset of the SEU data also appears as an appendix in Geluykens [1984]; since our study is narrower in scope than Geluykens',¹ we cite here a reduced version of his corpus.

The syntax of the SEU line-numbers is as follows. All numbers carry the prefix s; the first two numerical fields (e.g. 1.4) indicate subdivisions of the corpus. The final numerical field indicates the line number of the appropriate subdivision at which the cleft itself appears. Because our study is concerned with clefts in context, we also include a small quantity of preceding and subsequent material with each example. In the spoken data, the beginning of each cleft is indicated by ∇ , and the end by Δ .

In the second section, our own examples are given. Less context appears with these, because of the limitations of casual recording by means of pencil and paper.

Throughout both sections containing spoken data or references to such data, the notes [A] and [B] may appear after a reference or example. These notes denote sub-corpus A and sub-corpus B respectively, sub-corpora of data used and described in chapter 5 above.

¹Geluykens [1984] discusses headed cleft constructions beginning *All that ...*, *the thing that ...*, and *the one who* The study here does not consider these, as explained in chapter 1.

Spoken Data from the SEU Corpus

It-Clefts (54 examples)

Line Number	Sub-corpora	Line Number	Sub-corpora
S.1.1.57	[A]	S.2.8.09	[A] [B]
S.1.1.245	[A] [B]	S.2.8.901	[A] [B]
S.1.1.778	[A] [B]	S.2.9.1120	[A] [B]
S.1.2.90	[A] [B]	S.2.10.169	[A] [B]
S.1.2.329	[A] [B]	S.2.11.274	[A] [B]
S.1.2.1022	[A] [B]	S.2.11.699	[A] [B]
S.1.2.1336	[A] [B]	S.2.11.701	[A] [B]
S.1.5.401	[A] [B]	S.2.11.805	[A]
S.1.5.610	[A] [B]	S.2.11.1449	[A] [B]
S.1.8.1033	[A] [B]	S.2.11.1591	[A] [B]
S.1.9.349	[A]	S.2.12.1004	[A] [B]
S.1.10.83	[A] [B]	S.1.13.88	[A] [B]
S.1.10.125		S.2.13.265	
S.1.10.356	[A]	S.2.13.1190	[A] [B]
S.1.11.272	[A] [B]	S.2.13.1201	[A] [B]
S.1.12.1412	[A] [B]	S.2.14.880	[A] [B]
S.1.12.350		S.3.1.213	[A] [B]
S.1.13.581	[A] [B]	S.3.1.1123	[A] [B]
S.1.13.702	[A] [B]	S.3.1.1133	[A] [B]
S.1.13.854	[A] [B]	S.3.1.1356	[A] [B]
S.2.1.180	[A]	S.3.3.274	[A] [B]
S.2.1.1121	[A]	S.3.3.1151	[A]
S.2.2.661	[A] [B]	S.3.3.1154	[A] [B]
S.2.4.1367	[A] [B]	S.3.3.1256	[A] [B]
S.2.5.292	[A] [B]	S.3.4.863	[A]
S.2.5.625	[A] [B]	S.3.5.436	[A]
S.2.7.296	[A] [B]	S.3.6.583	

Wh-Clefts (63 examples)

Line Number	Sub-corpora	Line Number	Sub-corpora
S.1.1.13	[A] [B]	S.2.3.657	[A] [B]
S.1.2.80	[A] [B]	S.2.3.1037	[A] [B]
S.1.2.110	[A] [B]	S.2.4.137	[A] [B]
S.1.2.242	[A] [B]	S.2.4.727	[A] [B]
S.1.2.983	[A] [B]	S.2.5.94	[A] [B]
S.1.4.245	[A] [B]	S.2.7.213	[A] [B]
S.1.4.928	[A] [B]	S.2.8.158	[A]
S.1.5.207	[A] [B]	S.2.8.866	[A] [B]
S.1.5.879	[A] [B]	S.2.9.778	[A] [B]
S.1.5.1146	[A] [B]	S.2.9.943	[A] [B]
S.1.6.321	[A] [B]	S.2.9.1000	[A] [B]
S.1.6.603	[A]	S.2.10.185	[A] [B]
S.1.7.173	[A] [B]	S.2.10.1153	[A] [B]
S.1.7.625	[A]	S.2.11.153	[A]
S.1.7.761	[A] [B]	S.2.12.1138	[A] [B]
S.1.7.871	[A]	S.2.12.1357	
S.1.9.202	[A]	S.3.1.63	[A] [B]
S.1.10.827	[A] [B]	S.3.1.666	[A] [B]
S.1.11.760	[A] [B]	S.3.1.1354	
S.1.11.999	[A] [B]	S.3.2.802	
S.1.13.81	[A]	S.3.2.892	
S.1.13.148	[A] [B]	S.3.2.922	
S.1.13.878	[A] [B]	S.3.3.93	
S.1.13.1009	[A] [B]	S.3.3.561	
S.1.13.1054	[A] [B]	S.3.3.1239	
S.1.14.974	[A] [B]	S.3.4.47	
S.2.1.45	[A] [B]	S.3.4.154	
S.2.1.214	[A]	S.3.4.419	
S.2.2.1131	[A] [B]	S.3.4.548	
S.2.3.185	[A] [B]	S.3.5.1072	
S.2.3.367	[A]	S.3.6.268	
S.2.3.646	[A] [B]		

Reverse *Wh*-Clefts (106 examples)

Line Number	Sub-corpora	Line Number	Sub-corpora	Line Number	Sub-corpora
S.1.1.80	[A] [B]	S.1.13.130	[A] [B]	S.2.10.1350	
S.1.1.299	[A] [B]	S.1.13.568	[A] [B]	S.2.11.270	
S.1.1.1083	[A] [B]	S.1.13.958	[A] [B]	S.2.11.939	
S.1.2.35	[A] [B]	S.1.13.974	[A] [B]	S.2.11.978	
S.1.2.823	[A] [B]	S.1.14.165	[A] [B]	S.2.11.1574	
S.1.2.831	[A] [B]	S.2.2.749	[A] [B]	S.2.12.58	
S.1.3.361	[A] [B]	S.2.2.935	[A] [B]	S.2.12.110	
S.1.4.365	[A] [B]	S.2.2.1138	[A]	S.2.12.385	
S.1.4.881	[A] [B]	S.2.2.1150	[A] [B]	S.2.12.753	
S.1.5.10	[A] [B]	S.2.4.1131	[A] [B]	S.2.12.950	
S.1.5.377	[A] [B]	S.2.5.210	[A] [B]	S.2.13.165	
S.1.5.984		S.2.5.431	[A] [B]	S.2.13.847	
S.1.5.987	[A] [B]	S.2.5.620	[A] [B]	S.2.14.311	
S.1.6.17	[A] [B]	S.2.5.657		S.2.14.610	
S.1.6.107	[A] [B]	S.2.5.930	[A] [B]	S.2.14.641	
S.1.6.341		S.2.5.1106	[A] [B]	S.2.14.782	
S.1.6.781	[A] [B]	S.2.5.1124		S.2.14.806	
S.1.6.1131	[A] [B]	S.2.6.241	[A] [B]	S.2.14.957	
S.1.7.26	[A] [B]	S.2.6.400	[A] [B]	S.3.1.422	
S.1.7.464	[A] [B]	S.2.6.611		S.3.1.1410	
S.1.7.807	[A]	S.2.6.664		S.3.2.54	
S.1.8.321	[A] [B]	S.2.7.359		S.3.2.560	
S.1.8.768	[A] [B]	S.2.7.426		S.3.2.696	
S.1.8.790		S.2.8.13		S.3.2.875	
S.1.8.799	[A] [B]	S.2.8.17		S.3.2.1165	
S.1.9.560	[A] [B]	S.2.8.90		S.3.3.151	
S.1.9.623	[A] [B]	S.2.8.283		S.3.3.165	
S.1.9.651	[A] [B]	S.2.8.529		S.3.4.699	
S.1.9.867	[A] [B]	S.2.9.748		S.3.5.716	
S.1.9.1126	[A] [B]	S.2.9.874		S.3.5.734	
S.1.9.1412	[A] [B]	S.2.10.14		S.3.5.983	
S.1.10.447	[A] [B]	S.2.10.203		S.3.5.1012	
S.1.11.106	[A] [B]	S.2.10.382		S.3.6.440	
S.1.11.1060	[A] [B]	S.2.10.950		S.3.6.565	
S.1.12.57	[A] [B]	S.2.10.1180		S.3.6.833	
S.1.12.710	[A]	S.2.10.1321			

Author's Own Collection of Casual Conversational Data

It-Clefts (37 examples)

- (1) HP: We could find someone who's into Unix
SB: ∇It's oPentop that's the *problem*.△ I haven't been getting too much into Unix. [B]
- (2) HP: I was actually going to write and specify to the regions, but I sent it off anyway, 'cos ∇it was FRIday I was doing it△ [B]
- (3) RD: I thought you'd be upset about the cup.
JD: No, ∇it's the GLASS I'm *worried* about.△ [B]
- (4) JD: I'm really boring. Oops, I have to stop telling Jeremy I'm boring. Actually, I'm an incredibly exciting person.
RD: ∇It's ME that's *boring*.△
- (5) This has certainly been a family affair. Barbara Bush has been on all the TV networks, and ∇it was his Mexican *daughter* who seconded the nomination for the vice PRESIDency.△ So for a man who stated that he didn't want his family involved, this has been a bit of a U-turn.
- (6) JMHO: Are you sure nothing got copied across?
JD: (goes and looks) Yes, Jeremy, you're right. So ∇it's just the GARBage that didn't get copied.△ [B]
- (7) A: I want some cheese.
B: Which one did you want?
A: ∇It's CHEDDar I'm wanting△
- (8) JD: I think your Mum phoned. It was a middle-aged woman, anyway.
KK: (Rings up, talks for about 12 minutes)
JD: ∇Was it your Mum?△
- (9) JD: And here was me thinking that Robert had brought me the croissant, while all the time ∇it was YOU.△ [B]
- (10) JD: Did the subject really make this error?
PT: Sic? No, that's not sic. ∇It's just *me* that can't TYPE.△
- (11) JD: Well, we lost three surveys before we got our flat.
PT: And ∇it was on the FOURTH one you got it△
JD: Yeah [B]
- (12) PT: I've killed myself running. I can't go downstairs.
JD: (pointing to front of legs) ∇It's THESE muscles that *hurt*△, isn't it
PT: Yeah, and my FEET for some reason
- (13) JD: Fascist.
RD: ∇It's YOU that's the fascist△.
- (14) You see yesterday ∇it wasn't just US who put up their interest rates. It was the FRENCH, it was the SWISS, it was the GERMans△.

- (15) JD: These [oatcakes] always smell a bit rancid. ∇It's the OIL they use△, I think.
- (16) MM: The door's always wide open all weekend when the place is empty, but when there's a full department it gets locked
JD: From this we can conclude that ∇it's the people who aren't in at weekends that CLOSE it△
- (17) KV: I don't like going to popmobility. It's too energetic. ∇It's more the bendy, FLEXy stuff I need△. [B]
- (18) JD: But have I started to write mine? Have I? I've got four pages. ∇It's not the *start*ing, it's the carrying ON that's the *problem*△.
- (19) JD: Aha, THAT's what it is. Robert's doing a dump of brouwer.
AL: ∇Is it *Robert* that's making it go SLOW△?
- (20) JD: (pointing to a picture on the wall) It's a little cartoon, isn't it?
RD: Yes, ∇it's JON that did it△.
- (21) KK: The people realised that they needed a computer-∇it was an APPLE they needed△-and
- (22) HP: It gives you wordwrap when typing things. ∇It's as you say when you start INSERTing things that it gets to be a *problem*△.
- (23) JD: And does the head know?
HP: No. Oh, wait a minute. ∇It was the head who ARRANGED it△.
- (24) HP: Then she can decide which one of the rooms she wants, and I don't mind if ∇it's F10 or F13 I GET△
- (25) RD: Are you sure instant decaff is OK for you?
MC: Yes, ∇it's just the CAFFEine that I can't take *much* of△
- (26) JB: Isn't it locked? It's usually locked.
JD: ∇It's only the CUPboard that's *locked*△.
- (27) MM: So do you think they're going to wake up?
EK: So it's them and us, is it?
MM: Well, ∇it's me versus everyone else. It's people who are still working on their PhDs that we're *talking* about here△. [B]
- (28) JD: All you have to do is go to Leicestershire
MM: Or even Lancashire. ∇It was in Lancashire that I HEARD it△.
- (29) KK: In order that this doesn't get out of hand we have to be careful-keep it all in separate sections
HP: Well we can do the lower level specifications of what's needed separately. I mean ∇it's basically PROTOCOLs you're specifying△, that they should be able to follow
- (30) Telecom. ∇It's YOU we *answer* to△.
- (31) Who else would choose one of the gurus of the film industry, Orson Welles, to expound the philosophies that form the major insights of the film? To be fair, though, ∇it's Welles who expounds the most interesting ideas△.

- (32) FH: I've ordered.
 Waiter: 1, 2, 3, 4, 5 - ah
 RD: ∇It's ME who *hasn't* ordered△ [B]
- (33) A: We've found that the cabbage root fly comes in a wave motion. So by predicting when it's going to come, we can tell the farmer when to spray.
 B: But ∇it's not only *cabbages* that are *affected*, is it?△
 A: No, the cabbage root fly affects ALL the brassicas [B]
- (34) MM: You could try Elsa. She's got Chomsky's complete works
 CR: I know ... ∇It's cos of *her* that I NEED it△. Anyway, thanks
- (35) CM: When we set up the description space, ∇it's *these* kinds of relationship that we're interested in *computing*△. In particular, this one here, the relationship between two concepts of animacy.
- (36) CM: So my conclusion said I needed a *NECESSARY* condition. ∇It's sufficient ones I need△. Yes, you're right.
- (37) CM: ...but ∇it's essentially a very *simple* part of an *equation* that's being added *on*△.

Wh-Clefts (36 examples)

- (38) RD: Do you know what I'm doing? I'm testing my memory. ∇What it *does* is tests the whole of *memory* by *writing* to it and reading *back* from it to see if what it *wrote* is still *THERE*△. It does it over a thousand times.
- (39) RD: Come and have some Dairy Milk. ∇What *you* need is a wee bit of *CHOCOLATE*△. [B]
- (40) RD: ∇What you *could* do is hit it with a *HAMMER*△.
- (41) RD: They're OK, and if they're not, it's not clear what's wrong with them.
 JD: ∇What's *naff* about them is that they'd *normally* be w H clefts△.
- (42) RD: Is any of you guys shit-hot on *DAGS* and things of that sort?
 GM: No.
 RD: ∇What I want to *FIND* is a model of things I can use—a sort of formal description ... some formal thing I can use to follow the definitions of these feature structures I use△
- (43) MR: Do you want a recursive definition of a *DAG*, or something?
 RD: No, ∇what I want to *say* is that each *STRUCTURE* has the following characteristics△ ...
- (44) ∇Why we're *unhappy* about this scheme is that we don't believe it will *WORK*△.
- (45) They must have said to themselves, this man has gone crazy. He is mad. ∇What they *didn't* realise was that I was the one who was *sane*△.

- (46) HP: Well ∇ what you want is to have the word INSERTed, so that it simply goes in between the two Δ .
JD: But is that what we WANT?
- (47) I want them to understand who's in charge. And ∇ who's in *charge* is Neil KINNOck Δ .
- (48) HP: I mean there's already something of that in the low-level programming documentation
KK: ∇ What we actually NEED is a specification of what these things actually DO—rather than just what they *ought* to do Δ . [B]
- (49) HP: ∇ What we've been *working* on is the general framework for the evaluation Δ . And what I'm going to talk about is just briefly a framework for the pilot phase.
- (50) HP: @ What we've been working on is the general framework for the evaluation. And ∇ what I'm going to *talk* about is just briefly a framework for the Pilot phase Δ .
- (51) HP: The fifth section is a pragmatic one—a sort of dustbin category. ∇ What I *haven't* addressed is THIS Δ [points to chart] and what I haven't included in the framework is the question of how the project management is evaluated.
- (52) HP: The fifth section is a pragmatic one—a sort of dustbin category. What I haven't addressed is this [points to chart] and ∇ what I haven't included in the framework is the question of how the project MANAGEMENT is evaluated Δ .
- (53) JD: You seem to be able to do almost anything with them. ∇ What you *can't* do is write CHEQUES Δ .
- (54) LS: I've lost the application form, so ∇ what I'm trying to *do* is FIND it Δ . So what I'm doing now is going everywhere I've been, you know?
- (55) LS: I've lost the application form, so what I'm trying to do is find it. So ∇ what I'm doing *now* is going everywhere I've BEEN, you know Δ ?
- (56) A: So are you concerned about the allegations of safety on the rig?
B: Yes, of course I'm concerned, but ∇ what I'm concerned about is to get to the BOTTOM of them ... to get to the hard facts of what happened Δ . [B]
- (57) If you look at the figures, they are as I say: 101 accidents in 1986, but only 56 in 1987. ∇ What I'm concerned about is that there's been an improvement Δ .
- (58) CM: When I come to the last bit, ∇ what I'm trying to *do* is ... I'd just like to give a general flavour of what I'm trying to do Δ , but won't be going into very much detail.
- (59) CM: ∇ What I'm interested in *doing* is embodying general PROPERTIES like *this* one Δ .
- (60) CM: Well ∇ what you get is a description space, and we can think of a description space as being a lattice Δ .

- (61) CM: ∇What my project is *looking* at is certain kinds of encoding that we might be able to CONSTRUCT for description systemsΔ.
- (62) CM: So ∇what I want to *do* is just to tell you a RESULT that applies to finite tree encodable descriptionsΔ
- (63) CM: I want to be able to talk about elements of the description being prime. And ∇what *prime* means is that anything BELOW this description has that primitive *in* itΔ. [B]
- (64) CM: So in effect, ∇what this is *saying* is that any clash of GROUNDing has to be describable within *primitives*Δ. [B]
- (65) ∇What I want to *talk* about is basically a way of constructing BIG Is out of *smaller* onesΔ. And I need to have a way of talking about the bigger elements of these lattices.
- (66) CM: People have been using trees before, but in an ad hoc way. Partly ∇what I was trying to *do* was come upon a formal DESCRIPTION of what these things *meant*Δ.
- (67) CM: ...a conjunction of two addresses. And ∇what it's *pointing* to ...oops I missed a D out there ...is the thing pointed at by the FIRST addressΔ ...the conjunction of that and the thing pointed at by the SECond address. [B]
- (68) CM: If you just want to say look, these two descriptions simply aren't REATED to one another, ∇what you would *get* would be THISΔ [pointing to diagram].
- (69) CM: ...an operation I call deepening. ∇What you *do* is you take a description space D, with one *element*, and you say that where I had this element here before, I can replace that element with a whole description spaceΔ. So the result of that would be this lattice here.
- (70) CM: And here we have a tree. ∇How I *got* to the tree was simply by adding one *on* each timeΔ. What that's doing is essentially adding on a dimension each time. [B]
- (71) CM: How I got to the tree was simply by adding one on each time. ∇What that's *doing* is essentially adding on a DIMENSION each timeΔ.
- (72) JD: I quite like documentaries. ∇What I loathe are those interminable heart transplantsΔ. [B]

Reverse Wh-Clefts (16 examples)

- (73) JD: Why don't you use a spanner to open it? Or do you need something else?
RD: ∇A SPANNER's what I *need*Δ
- (74) MG: It must be Lawrence Erlbaum asking him to write a book, and he's got to say, 'I'm sorry, that's not the plan. I've got to write the article for the Manchester Guardian first.'
JD: ∇Is that what he WANTEDΔ?

- (75) JD: Aha, ∇ THAT's what it is Δ . Robert's doing a dump of brouwer.
AL: Is it *Robert* that's making it go SLOW? [A] [B]
- (76) SB: If you drop it in here it puts it in here.
HP: ∇ That's obviously what you WANTED when you pointed there Δ .
- (77) HP: Well what you want is to have the word INSERTED, so that it simply goes in between the two.
JD: But ∇ is that what we WANT Δ ? [B]
- (78) JD: ∇ I didn't know THAT was where he lived Δ .
- (79) JD: That was my friend Mike.
MM: Gosh, ∇ THAT's what he looks like Δ . [A] [B]
- (80) MM: I thought you were going to Spain.
JD: Yes, ∇ that's what I told Mimo Δ , but ...
- (81) MM: But it can't be true that they're simply to make intonation clearer, as then you don't have account for written language
JD: Yes, ∇ THAT's what you would *think* Δ , but you can make a case for a model of reading that uses intonational cues.
- (82) RD: I'm phoning to complain ...sent me a letter that accuses me, because ∇ that's what it DOES Δ , of failing to either return the letter or specify enough information.
- (83) JD: How are you with things like cabbage and cauliflower?
MC: ∇ Vegetables are what I should eat MOST of Δ , so they're fine.
- (84) JD: They call it a 'Bain Marie', which means 'Mary's bath'. I don't know why. Maybe ∇ she was who INVENTED it Δ .
- (85) JD: I've been given a really nice Habitat holepunch for Christmas. It's matt black
MM: Can't you go and exchange it for a stapler? Because ∇ THAT's what I *really* need Δ
- (86) JD: We're not having a very good relationship at the moment, because he's on a diet. He's being a right pain. Robert practically threw him at the wall last night. [To cat] ∇ You were what woke us up at four o'CLOCK Δ , weren't you?
- (87) A: So are you concerned about the allegations of safety on the rig?
B: Yes, of course I'm concerned, but what I'm concerned about is to get to the bottom of them ...to get to the hard facts of what happened. And ∇ THAT's what I'm determined to *do* Δ .
- (88) FH: ...or are these just what you could think of at the TIME?
CM: ∇ They're what I could THINK of Δ ...I'd like to think of MORE